

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:11:52 ; Search time 93.4211 Seconds
(without alignments)
47.032 Million cell updates/sec

Title: US-10-774-242A-4
Perfect score: 65
Sequence: 1 YALMWSNGK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : A_Geneseq 21:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*
9: Geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	10	ADR47260	Human alp
2	65	100.0	51	AAM17490	Peptide #
3	65	100.0	51	ABB36511	Peptide #
4	65	100.0	51	AAM30010	Peptide #
5	65	100.0	51	AAM69678	Human bon
6	65	100.0	51	ABG51357	Human liv
7	65	100.0	51	ABG39294	Human pep
8	65	100.0	57	ABO55658	Human gen
9	65	100.0	723	AEb94227	Human sol
10	65	100.0	734	AEb94218	Human sol
11	65	100.0	750	AEb94161	Human sol
12	65	100.0	759	AAW31963	Human fib
13	65	100.0	760	AAW27438	Human fib
14	65	100.0	760	ABr47452	Breast ca
15	65	100.0	760	ADN95552	Human BEC
16	65	100.0	760	ADQ21351	Human sof
17	65	100.0	760	ADW14775	Tumor-ass
18	65	100.0	760	AEb94159	Human wil
19	60	92.3	761	AEb94163	Mouse wil
20	55	84.6	504	ADH17327	Polypepti
21	53	81.5	493	ABP55609	Human Dpp
22	53	81.5	593	AAR40916	Sequence
23	53	81.5	593	AAR54614	Delta594-
24	53	81.5	688	ADO71642	Amino aci

25	53	81.5	736	8	ADO40240	Human Dpp
26	53	81.5	739	2	AAR54613	Delta24-3
27	53	81.5	759	2	AAR54612	Delta3-9
28	53	81.5	766	2	AAR40909	Sequence
29	53	81.5	766	2	AAR54611	Native CD
30	53	81.5	766	5	ABB08991	Human dip
31	53	81.5	766	5	ABG61910	Prostate
32	53	81.5	766	5	AAG78417	Human dip
33	53	81.5	766	5	AAO15555	Human dip
34	53	81.5	766	6	ABP55629	Human dpp
35	53	81.5	766	6	ABP56700	Human liv
36	53	81.5	766	7	ADD14045	Human arc
37	53	81.5	766	7	ADD27855	Human dip
38	53	81.5	766	7	ADD46934	Human dip
39	53	81.5	766	7	ADN39272	Cancer/an
40	53	81.5	766	7	ADN39604	Cancer/an
41	53	81.5	766	8	ADJ83981	Human ful
42	53	81.5	766	8	ADJ75313	Marker ge
43	53	81.5	766	8	ADO19398	Human PRO
44	53	81.5	766	8	ADO19400	Human PRO
45	53	81.5	766	8	ADO19806	Human PRO
46	53	81.5	766	8	ADO19806	Human PRO
47	53	81.5	766	8	ADO71644	Amino aci
48	53	81.5	766	8	ADO80365	Dipeptidy
49	53	81.5	766	8	ABM80355	Tumour-as
50	53	81.5	766	8	ADP54458	Human PRO
51	53	81.5	766	8	ADU06688	Novel bro
52	53	81.5	766	8	ADV25525	Human dip
53	53	81.5	766	8	ADY15161	PRO poly
54	53	81.5	766	9	ADY16580	PRO poly
55	53	81.5	766	9	ADY14038	Human dip
56	53	81.5	766	9	ABE77579	Human dip
57	53	81.5	766	9	ABE94223	CD26/dipe
58	53	81.5	767	3	AB11748	Rat dipep
59	53	81.5	767	6	ABP56699	Rat liver
60	53	81.5	767	7	AD446932	Rat prote
61	53	81.5	767	7	ABE77580	Rat dipep
62	47	72.3	760	8	ADJ76138	Marker ge
63	47	72.3	760	8	ADO71646	Amino aci
64	47	72.3	760	9	ABE77581	Mouse dip
65	47	72.3	760	9	ABE94226	Mouse CD2
66	47	72.3	829	8	ADN24353	Bacterial
67	46	70.8	771	2	AAW89589	Aspergill
68	46	70.8	771	2	AAW97798	Aspergill
69	45	69.2	574	9	ABM92606	M. xanthu
70	45	69.2	661	8	ADS30442	Bacterial
71	45	69.2	758	8	ADN19648	Bacterial
72	44	67.7	323	7	ADG89853	Haemophil
73	43	66.2	112	4	ABG09950	Novel hum
74	43	66.2	189	5	AAU79255	Human mem
75	43	66.2	189	5	AAU79257	Rat membr
76	43	66.2	189	5	AAU79258	Rabbit me
77	43	66.2	189	5	AAU79256	Murine me
78	43	66.2	222	9	ABE37458	L. pneumo
79	43	66.2	222	9	ABE40772	L. pneumo
80	43	66.2	498	3	AAQ49364	Arabidops
81	43	66.2	498	3	AAQ17973	Arabidops
82	43	66.2	549	8	ADP29329	Human sec
83	43	66.2	579	2	ABR86406	Human mat
84	43	66.2	581	2	ABE94410	Human mat
85	43	66.2	582	2	AAR86407	Human mat
86	43	66.2	582	2	AAW52134	Human pia
87	43	66.2	582	2	AAW52134	Rabbit me
88	43	66.2	582	2	ABR84616	Amino aci
89	43	66.2	582	4	AAE10423	Human mat
90	43	66.2	582	5	AAU84294	Human end
91	43	66.2	582	5	AAE21037	Human mem
92	43	66.2	582	5	AAW50865	Matrix me
93	43	66.2	582	7	ADN15498	Human bas
94	43	66.2	582	7	ADN64179	Human PRO
95	43	66.2	582	7	ADN64177	Rat Prote
96	43	66.2	582	7	ADN13708	Tumor-Asp
97	43	66.2	582	7	ADN95889	Human BEC

98	43	66.2	582	8	ADN07703	Adn07703 Human mat	171	63.1	1113	7	ADG31718	Human pro	
99	43	66.2	582	8	ABM81541	ABM81541 Tumour-as	172	41	1132	6	ADA54352	Human pro	
100	43	66.2	582	8	ADP23426	ADP23426 PRO polyo	173	41	1352	7	ADG31722	Human pro	
101	43	66.2	582	9	ABE94378	ABE94378 Human Mtl	174	41	3065	6	ABP96236	Human nuc	
102	43	66.2	609	3	AAG49363	Aag49363 Arabidops	175	40	61.5	18	ADW28772	Core homo	
103	43	66.2	609	3	AAG17972	Aag17972 Arabidops	176	40	61.5	18	ADX97371	BMP conta	
104	43	66.2	630	3	AAG17971	Aag17971 Arabidops	177	40	61.5	19	ADW28773	Core homo	
105	43	66.2	630	3	AAG49362	Aag49362 Arabidops	178	40	61.5	19	ADX97372	Rat scler	
106	42	64.6	48	6	ABP55608	Abp55608 Human DPP	179	40	61.5	24	ADW28795	Mouse LRP	
107	42	64.6	336	2	AAW86338	AAW86338 Kidney in	180	40	61.5	26	ADW28795	Mouse LRP	
108	42	64.6	352	2	AAW86338	AAW86338 Kidney in	181	40	61.5	26	ADW28795	Immunoglo	
109	42	64.6	380	6	ABP55572	Abp55572 Full leng	182	40	61.5	27	ADW28795	Mouse LRP	
110	42	64.6	400	7	ADP05536	Adp05536 Bacterial	183	40	61.5	52	ADI27219	Mouse LRP	
111	42	64.6	438	6	ABU29828	Abu29828 Protein e	184	40	61.5	178	ABM91764	M. xanthu	
112	42	64.6	443	7	ADN97492	Adn97492 E. faeciu	185	40	61.5	178	ABM91764	M. xanthu	
113	42	64.6	579	8	ADN20537	Adn20537 Bacterial	186	40	61.5	188	ADW28839	Murine TG	
114	42	64.6	664	8	ADN25524	Adn25524 Bacterial	187	40	61.5	188	ADW28839	Murine TG	
115	42	64.6	691	5	ABG61612	Abg61612 Human DPR	188	40	61.5	190	ADW28761	Rat TGF b	
116	42	64.6	706	5	ABG61611	Abg61611 Human DPR	189	40	61.5	190	ADW28761	Rat scler	
117	42	64.6	743	5	ADR43716	Adr43716 Human pro	190	40	61.5	205	ADX97360	Mouse LRP	
118	42	64.6	746	6	ABP55582	Abp55582 Human DPP	191	40	61.5	211	ADW27148	Murine TG	
119	42	64.6	746	6	ABP55584	Abp55584 Human DPP	192	40	61.5	211	AAW96432	Mouse clo	
120	42	64.6	746	6	ABP55581	Abp55581 Human DPP	193	40	61.5	211	ABM07210	Mouse clo	
121	42	64.6	775	9	ADY51819	Ady51819 T. rubrum	194	40	61.5	211	ADW28839	Murine TG	
122	42	64.6	789	5	ABP43687	Abp43687 Dipteryd	195	40	61.5	211	ADW28839	Murine TG	
123	42	64.6	789	6	ABP55583	Abp55583 Human DPP	196	40	61.5	211	ADW28761	Rat scler	
124	42	64.6	789	6	ABP55577	Abp55577 Mouse DPP	197	40	61.5	211	ADW27147	Mouse LRP	
125	42	64.6	796	5	ABG61593	Abg61593 Human DPP	198	40	61.5	211	ADW28708	Murine TG	
126	42	64.6	796	5	ABG61593	Abg61593 Human DPP	199	40	61.5	211	ADW28708	Murine TG	
127	42	64.6	796	5	ABG61593	Abg61593 Human DPP	200	40	61.5	211	ADW28708	Murine TG	
128	42	64.6	796	6	ABP55592	Abp55592 DPP10 pro	201	40	61.5	213	ADW28708	Murine TG	
129	42	64.6	796	6	ABP55576	Abp55576 Mouse DPP	202	40	61.5	213	ADW28708	Murine TG	
130	42	64.6	796	6	ABP55591	Abp55591 DPP10 tra	203	40	61.5	213	ADW28708	Murine TG	
131	42	64.6	796	6	ABP55624	Abp55624 Human DPP	204	40	61.5	213	ADW28708	Murine TG	
132	42	64.6	796	6	ABP55574	Abp55574 Mouse DPP	205	40	61.5	213	ADW28708	Murine TG	
133	42	64.6	796	6	ABP55580	Abp55580 Human DPP	206	40	61.5	213	ADW28708	Murine TG	
134	42	64.6	796	6	ABP55628	Abp55628 Human DPP	207	40	61.5	213	ADW28708	Murine TG	
135	42	64.6	796	6	ABP55625	Abp55625 Mouse DPP	208	40	61.5	213	ADW28708	Murine TG	
136	42	64.6	796	7	ADA09104	Ada09104 Novel hum	209	40	61.5	213	ADW28708	Murine TG	
137	42	64.6	797	6	ABP55575	Abp55575 Mouse DPP	210	40	61.5	213	ADW28708	Murine TG	
138	42	64.6	797	6	ABP55573	Abp55573 Human DPP	211	40	61.5	213	ADW28708	Murine TG	
139	42	64.6	798	7	ADE47758	AdE47758 Human NOV	212	40	61.5	213	ADW28708	Murine TG	
140	42	64.6	798	8	ADJ79028	Adj79028 Human NOV	213	40	61.5	213	ADW28708	Murine TG	
141	42	64.6	799	6	ABP55578	Abp55578 Mouse DPP	214	40	61.5	213	ADW28708	Murine TG	
142	42	64.6	800	6	ABP55579	Abp55579 Mouse DPP	215	40	61.5	213	ADW28708	Murine TG	
143	42	64.6	818	8	ADS43599	Ads43599 Bacterial	216	40	61.5	213	ADW28708	Murine TG	
144	42	64.6	960	3	AAG48128	Aag48128 Arabidops	217	40	61.5	213	ADW28708	Murine TG	
145	42	64.6	968	3	AAG48127	Aag48127 Arabidops	218	40	61.5	213	ADW28708	Murine TG	
146	42	64.6	1063	3	AAG48126	Aag48126 Arabidops	219	40	61.5	213	ADW28708	Murine TG	
147	41	63.1	63	8	ABG55912	Abg55912 Human gen	220	40	61.5	213	ADW28708	Murine TG	
148	41	63.1	69	8	ADK01476	Adk01476 Hepatitis	221	40	61.5	213	ADW28708	Murine TG	
149	41	63.1	199	4	AAAB61488	AAAB61488 Murine MA	222	40	61.5	213	ADW28708	Murine TG	
150	41	63.1	257	8	ADR86370	Adr86370 Aspergill	223	40	61.5	213	ADW28708	Murine TG	
151	41	63.1	365	9	ADY65142	Ady65142 S. mansoni	224	40	61.5	213	ADW28708	Murine TG	
152	41	63.1	458	5	ABP73738	Abp73738 Candida a	225	40	61.5	213	ADW28708	Murine TG	
153	41	63.1	530	8	ABG68894	Abg68894 Plant ful	226	40	61.5	213	ADW28708	Murine TG	
154	41	63.1	557	8	ADL57099	Adl57099 Human NOV	227	40	61.5	213	ADW28708	Murine TG	
155	41	63.1	657	2	AAW71851	Aaw71851 Mouse mem	228	40	61.5	213	ADW28708	Murine TG	
156	41	63.1	669	2	AAW71851	Aaw71851 Mouse mem	229	40	61.5	213	ADW28708	Murine TG	
157	41	63.1	669	4	AAW71851	Aaw71851 Mouse mem	230	40	61.5	213	ADW28708	Murine TG	
158	41	63.1	669	4	AAW71851	Aaw71851 Mouse mem	231	40	61.5	213	ADW28708	Murine TG	
159	41	63.1	803	6	ABR58662	AbR58662 Human can	232	40	61.5	213	ADW28708	Murine TG	
160	41	63.1	803	7	ABW79818	Abw79818 Rat dipep	233	40	61.5	213	ADW28708	Murine TG	
161	41	63.1	804	6	ABP55627	Abp55627 Human dpp	234	40	61.5	213	ADW28708	Murine TG	
162	41	63.1	804	9	ADW26329	Adw26329 Novel cel	235	40	61.5	213	ADW28708	Murine TG	
163	41	63.1	859	7	ADE58039	AdE58039 Rat Prote	236	39	60.0	75	ABO72403	Human gen	
164	41	63.1	859	7	ADE58035	AdE58035 Rat Prote	237	39	60.0	81	ABO60100	Human tra	
165	41	63.1	859	9	ADW26403	Adw26403 Novel cel	238	39	60.0	82	ABP35151	Human tra	
166	41	63.1	865	6	ABP55626	Abp55626 Human dpp	239	39	60.0	154	3	ABG21528	Arabidops
167	41	63.1	865	7	ADE58041	AdE58041 Human Pro	240	39	60.0	164	3	ABG35926	Zea mayis
168	41	63.1	865	7	ADE58037	AdE58037 Human Pro	241	39	60.0	174	3	ABO8458	Amino aci
169	41	63.1	865	9	ADW26259	Adw26259 Novel cel	242	39	60.0	177	3	ABG35925	Zea mayis
170	41	63.1	964	9	ABE27316	AbE27316 Pinus rad	243	39	60.0	202	3	ABG35924	Zea mayis

244	39	60.0	214	3	AAG39988	Aag39988 Arabidops	317	38	58.5	381	5	ABP69594	Abp69594 Human pol
245	39	60.0	246	8	ADY11440	Ady11440 Plant ful	318	38	58.5	392	6	AAO26246	Aao26246 MDDT rela
246	39	60.0	250	4	AG000027	Ag000027 Novel hum	319	38	58.5	392	6	ABR58440	AbR58440 Human NOV
247	39	60.0	254	6	ABM67034	Abm67034 Photorhab	320	38	58.5	392	8	ADQ96100	Adq96100 T cell ac
248	39	60.0	254	6	ADY066418	Ady066418 Plant ful	321	38	58.5	399	5	ABP68984	Abp68984 Human pol
249	39	60.0	255	3	AB080471	Ab080471 Amino aci	322	38	58.5	399	6	ABP70710	Abp70710 Human lon
250	39	60.0	261	3	AG215127	Ag215127 Arabidops	323	38	58.5	413	6	ADA34302	Ada34302 Acinetoba
251	39	60.0	305	2	AAW71479	Aaw71479 Helicobac	324	38	58.5	417	7	ABO77124	AbO77124 Pseudomon
252	39	60.0	321	3	AG39987	Aag39987 Arabidops	325	38	58.5	441	6	ABP75500	Abp75500 Human sec
253	39	60.0	321	8	ADM56123	Adm56123 Thale cre	326	38	58.5	441	6	ADX92604	Adx92604 Plant ful
254	39	60.0	321	8	ADN72317	Adn72317 Thale cre	327	38	58.5	479	8	ADT60887	Adt60887 Plant pol
255	39	60.0	332	8	ADJ61773	Adj61773 Triticum	328	38	58.5	484	4	AAB68282	Aab68282 Amino aci
256	39	60.0	333	8	ADJ61769	Adj61769 Zea mays	329	38	58.5	485	4	AAB68284	Aab68284 Amino aci
257	39	60.0	333	8	ADJ61767	Adj61767 Hordeum v	330	38	58.5	485	4	AAB68516	Aab68516 Human GTP
258	39	60.0	360	8	ADX96935	Adx96935 Plant ful	331	38	58.5	485	4	AAB92844	Aab92844 Human pro
259	39	60.0	366	9	ABE27309	Aeb27309 Pinus rad	332	38	58.5	485	5	ABM97306	Abm97306 Novel hum
260	39	60.0	371	8	ADX68648	Adx68648 Plant ful	333	38	58.5	485	5	ABM81862	Abm81862 Tumour-as
261	39	60.0	399	8	ADN19222	Adn19222 Bacterial	334	38	58.5	485	8	ADP54460	Adp54460 Human PRO
262	39	60.0	418	9	ADW17547	Adw17547 Pinus rad	335	38	58.5	485	9	ADX07078	Adx07078 Cyclin-de
263	39	60.0	419	9	ABE40069	Aeb40069 L. pneumo	336	38	58.5	507	8	ADU47103	Adu47103 Human ery
264	39	60.0	423	9	ABE36677	Aeb36677 L. pneumo	337	38	58.5	711	8	ADY25145	Ady25145 Plant ful
265	39	60.0	443	7	ADD44907	Add44907 Rat Prote	338	38	58.5	745	4	ABB65409	Abb65409 Drosophi
266	39	60.0	465	6	ABR53417	AbR53417 Protein s	339	38	58.5	842	8	ADQ08602	Adq08602 Ciona int
267	39	60.0	465	7	ADK64696	Adk64696 Disease t	340	38	58.5	1278	5	ABP65361	Abp65361 Bifidobac
268	39	60.0	465	8	ADS44113	Ads44113 Bacterial	341	38	58.5	2129	4	AAU48629	Aau48629 Propionib
269	39	60.0	657	8	ADS30316	Ads30316 Bacterial	342	38	58.5	2129	6	ABM45148	Abm45148 Mycobacte
270	39	60.0	657	8	ADQ39692	Adq39692 Human myo	343	37	56.9	207	5	AAO15874	Aao15874 Mycobacte
271	39	60.0	655	9	ABE38091	Aeb38091 L. pneumo	344	37	56.9	10	4	AAB71052	Aab71052 Cellulose
272	39	60.0	733	8	ADP09480	Adp09480 Human pro	345	37	56.9	13	4	AAB70985	Aab70985 Cellulose
273	39	60.0	745	5	ADP65572	Adp65572 Bifidobac	346	37	56.9	13	4	AAB70984	Aab70984 Cellulose
274	39	60.0	900	8	ADN21103	Adn21103 Bacterial	347	37	56.9	13	4	AAB70986	Aab70986 Cellulose
275	39	60.0	981	9	ABM97613	Abm97613 M. xanthu	348	37	56.9	15	8	ADN17052	Adn17052 Gi-alpha
276	39	60.0	1070	7	ADQ94089	Adq94089 E. faeciu	349	37	56.9	64	6	AAU54897	Aau54897 Propionib
277	39	60.0	1072	5	ABE48188	AbE48188 Listeria	350	37	56.9	64	6	ABM51416	Abm51416 Propionib
278	39	60.0	1073	7	ADH88471	Adh88471 Enterococ	351	37	56.9	73	4	ABG04412	Abg04412 Novel hum
279	39	60.0	1234	8	ADQ39694	Adq39694 Human myo	352	37	56.9	79	4	ABG27411	Abg27411 Novel hum
280	39	60.0	1365	8	ADQ39691	Adq39691 Human myo	353	37	56.9	93	4	AAU29518	Aau29518 Novel hum
281	39	60.0	1366	8	ADQ39692	Adq39692 Human myo	354	37	56.9	93	4	AAU29518	Aau29518 Novel hum
282	39	60.0	1411	8	ADQ39693	Adq39693 Human myo	355	37	56.9	100	4	AAU62322	Aau62322 Propionib
283	39	60.0	2113	2	AAW00384	Aaw00384 Plasmodiu	356	37	56.9	100	6	ABM58841	Abm58841 Propionib
284	38	58.5	8	ADU08131	Adu08131 Heat shoc	357	37	56.9	101	3	ABG18979	Abg18979 Zea mays	
285	38	58.5	62	5	ADG62100	Adg62100 Human pro	358	37	56.9	112	3	AAAG1071	AaG1071 Arabidops
286	38	58.5	63	4	ABG07657	Abg07657 Novel hum	359	37	56.9	115	3	AAAG18978	AaG18978 Zea mays
287	38	58.5	66	4	ABU61787	Abu61787 Propionib	360	37	56.9	129	9	ABE63948	Aeb63948 Lysozyme
288	38	58.5	66	6	ABM58306	Abm58306 Propionib	361	37	56.9	129	9	ABE63949	Aeb63949 Lysozyme
289	38	58.5	68	4	AAU43329	Aau43329 Propionib	362	37	56.9	133	4	AAO03686	Aao03686 Human pol
290	38	58.5	68	6	ABM40048	Abm40048 Propionib	363	37	56.9	149	8	ADY72289	Ady72289 Plant ful
291	38	58.5	104	7	ADB64408	Adb64408 Human pro	364	37	56.9	149	8	ADY72289	Ady72289 Plant ful
292	38	58.5	146	6	ABP75894	Abp75894 Human sec	365	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
293	38	58.5	185	4	ABG12294	Abg12294 Novel hum	366	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
294	38	58.5	200	8	ADQ96280	Adq96280 T cell ac	367	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
295	38	58.5	200	8	ADQ96278	Adq96278 T cell ac	368	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
296	38	58.5	210	7	ABO75153	AbO75153 Pseudomon	369	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
297	38	58.5	220	3	AG39682	Aag39682 Arabidops	370	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
298	38	58.5	230	6	ABU41630	Abu41630 Protein e	371	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
299	38	58.5	232	6	ABM68707	Abm68707 Photorhab	372	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
300	38	58.5	232	6	ABU38897	Abu38897 Protein e	373	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
301	38	58.5	233	3	AG39681	Aag39681 Arabidops	374	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
302	38	58.5	237	3	ADF05265	Adf05265 Bacterial	375	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
303	38	58.5	240	8	AD341806	Ad341806 Bacterial	376	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
304	38	58.5	242	7	ABO65616	AbO65616 Klebsiell	377	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
305	38	58.5	248	4	ABG17902	Abg17902 Novel hum	378	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
306	38	58.5	255	4	ABG17728	Abg17728 Novel hum	379	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
307	38	58.5	258	3	AGS1378	Ags1378 Arabidops	380	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
308	38	58.5	261	4	AAB68283	Aab68283 Amino aci	381	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
309	38	58.5	311	4	AAU00782	Aau00782 Human apo	382	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
310	38	58.5	325	3	AGS1377	Ags1377 Arabidops	383	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
311	38	58.5	325	3	AGS1376	Ags1376 Arabidops	384	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
312	38	58.5	332	7	ABO82696	AbO82696 Pseudomon	385	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
313	38	58.5	340	6	ABU25979	Abu25979 Aspergill	386	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
314	38	58.5	354	6	ADQ96274	Adq96274 T cell ac	387	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
315	38	58.5	362	5	ABB07569	AbB07569 Human tum	388	37	56.9	151	7	ABO62592	AbO62592 Klebsiell
316	38	58.5	367	7	ABO67882	AbO67882 Pseudomon	389	37	56.9	151	7	ABO62592	AbO62592 Klebsiell

390	37	56.9	427	4	ABB71359	Abb71359 Drosophil	463	36	55.4	83	5	ABG61744	Abg61744 Novel ova
391	37	56.9	428	8	ADJ75330	Adj75330 Marker ge	464	36	55.4	84	6	ADA57060	Ada57060 Human sec
392	37	56.9	438	3	AAy44622	Aay44622 Truncated	465	36	55.4	84	7	ADC74216	Adc74216 Human sec
393	37	56.9	438	3	AAy44623	Aay44623 R154C tru	466	36	55.4	84	7	ADJ37925	Adj37925 Human sec
394	37	56.9	450	4	AAU56552	Aau56552 Propionib	467	36	55.4	85	2	AAy36306	Aay36306 Human sec
395	37	56.9	450	6	ABM53071	Abm53071 Propionib	468	36	55.4	85	2	ADa11655	Ada11655 Human nov
396	37	56.9	452	8	ADY13719	Ady13719 Plant ful	469	36	55.4	86	8	ADL92540	Adl92540 iMab1301
397	37	56.9	458	7	ADe28675	Ade28675 Human NOV	470	36	55.4	86	8	ADL1852	Adl1852 Versatile
398	37	56.9	458	8	ADM93420	Adm93420 Human NOV	471	36	55.4	86	9	ADx70168	Adx70168 Versatile
399	37	56.9	464	8	ADU02850	Adu02850 Novel hum	472	36	55.4	91	3	AGQ00079	Agq00079 Human sec
400	37	56.9	477	8	ADY04901	Ady04901 Plant ful	473	36	55.4	93	2	AAr72860	Aar72860 Immunoglo
401	37	56.9	492	5	AAm47905	Aam47905 Human fuc	474	36	55.4	96	4	AAU54486	Aau54486 Propionib
402	37	56.9	507	2	AAr06511	Aar06511 EPO recep	475	36	55.4	96	6	ABM51005	Abm51005 Propionib
403	37	56.9	507	2	AAr47517	Aar47517 MEL EPO r	476	36	55.4	99	9	ADL70353	Adl70353 Versatile
404	37	56.9	507	2	AAr50327	Aar50327 Mouse sol	477	36	55.4	101	8	ADL92538	Adl92538 iMab1300
405	37	56.9	507	2	AAr69502	Aar69502 Mouse ery	478	36	55.4	101	8	ADL1851	Adl1851 Versatile
406	37	56.9	508	2	AAr06512	Aar06512 EPO recep	479	36	55.4	101	9	ADx70186	Adx70186 Versatile
407	37	56.9	508	2	AAr47518	Aar47518 Human EPO	480	36	55.4	103	9	ADx70381	Adx70381 Versatile
408	37	56.9	508	2	AAr70032	Aar70032 Human ery	481	36	55.4	106	7	ADf04860	Adf04860 Bacterial
409	37	56.9	508	2	AAr69503	Aar69503 Human ery	482	36	55.4	106	8	ADY22446	Ady22446 Plant ful
410	37	56.9	508	5	ABM09173	Abm09173 Human ery	483	36	55.4	107	8	ADr42014	Adr42014 Hair and
411	37	56.9	508	7	ADe28677	Ade28677 Human NOV	484	36	55.4	116	4	ABG26368	Abg26368 Novel hum
412	37	56.9	508	8	ADM93422	Adm93422 Human NOV	485	36	55.4	129	9	ADV67234	Adv67234 Amino aci
413	37	56.9	508	8	ADQ05724	Adq05724 Human ery	486	36	55.4	139	5	ABP52943	Abp52943 Human lun
414	37	56.9	508	8	ADP12417	Adp12417 Protein e	487	36	55.4	171	4	ABG22740	Abg22740 Novel hum
415	37	56.9	508	8	ABM82037	Abm82037 Tumour-as	488	36	55.4	181	7	ABR52927	Abbr52927 Protein e
416	37	56.9	525	3	AAg31643	Aag31643 Arabidops	489	36	55.4	181	7	ADK62506	Adk62506 Disease t
417	37	56.9	539	6	ABU28254	Abu28254 Protein e	490	36	55.4	195	3	AGQ09786	Agq09786 Arabidops
418	37	56.9	540	6	ABU32404	Abu32404 Protein e	491	36	55.4	209	4	AAU65817	Aau65817 Propionib
419	37	56.9	544	6	ABM69644	Abm69644 Photorhab	492	36	55.4	209	6	ABM62336	Abm62336 Propionib
420	37	56.9	546	7	ADC31278	Adc31278 Human nov	493	36	55.4	213	3	AGQ09785	Agq09785 Arabidops
421	37	56.9	567	7	ABO64323	AbO64323 Klebsiell	494	36	55.4	213	4	AAU18651	Aau18651 Renal and
422	37	56.9	578	3	AAg31642	Aag31642 Arabidops	495	36	55.4	213	6	ABU97266	Abu97266 Human pol
423	37	56.9	596	5	AAU77251	Aau77251 Arabidops	496	36	55.4	215	8	ADx80619	Adx80619 Arabidops
424	37	56.9	597	3	AAg31641	Aag31641 Protein e	497	36	55.4	222	3	AAQ09784	Aaq09784 Arabidops
425	37	56.9	597	6	ABP81227	Abp81227 Arabidops	498	36	55.4	222	3	AAQ14451	Aaq14451 Arabidops
426	37	56.9	607	9	ABE27220	AbE27220 Pinus rad	499	36	55.4	226	8	ADN26835	Adn26835 Bacterial
427	37	56.9	635	8	ADY05348	Ady05348 Plant ful	500	36	55.4	231	6	ABU40218	Abu40218 Protein e
428	37	56.9	747	4	ABM67528	Abm67528 Drosophil							
429	37	56.9	762	3	AAb25146	Abb25146 Pinus rad							
430	37	56.9	774	2	AAy37454	Aay37454 Protein i							
431	37	56.9	779	3	AAb18787	Aab18787 Protein e							
432	37	56.9	779	8	ADN23970	Adn23970 Bacterial							
433	37	56.9	785	4	ABM61535	Abm61535 Drosophil							
434	37	56.9	799	8	ADN23969	Adn23969 Bacterial							
435	37	56.9	815	5	ABR04869	Abbr04869 Herbicida							
436	37	56.9	834	2	AAr04869	Aar04869 Alpha-i-6							
437	37	56.9	897	4	ABM66323	Abm66323 Drosophil							
438	37	56.9	931	2	AAW89596	Aaw89596 Saccharom							
439	37	56.9	931	8	ADs43967	AdS43967 Bacterial							
440	37	56.9	1101	4	AAU62902	Aau62902 Propionib							
441	37	56.9	1101	6	ABM59421	Abm59421 Propionib							
442	37	56.9	1422	3	AAg48628	Aag48628 Arabidops							
443	37	56.9	1430	3	AAg48627	Aag48627 Arabidops							
444	37	56.9	1487	3	AAg48626	Aag48626 Arabidops							
445	37	56.9	1534	8	ADN72343	Adn72343 Thale cre							
446	37	56.9	1534	9	ADZ44902	AdZ44902 A thalian							
447	37	56.9	2228	2	AAW93944	Aaw93944 P. fatcip							
448	37	56.9	13	9	ADY99970	Ady99970 Rat neu							
449	37	56.9	30	4	ABG17608	Abg17608 Novel hum							
450	37	56.9	47	4	ABG00996	Abg00996 Novel hum							
451	37	56.9	57	8	ABO56784	AbO56784 Human gen							
452	37	56.9	70	4	ABB41076	Abb41076 Peptide #							
453	37	56.9	70	4	AAm34852	Aam34852 Peptide #							
454	37	56.9	70	4	ABB25138	Abb25138 Protein #							
455	37	56.9	70	4	ABM74736	Abm74736 Human bon							
456	37	56.9	70	4	AAm61934	Aam61934 Human bra							
457	37	56.9	70	4	ABG56519	Abg56519 Human liv							
458	37	56.9	79	3	AAU54300	Aau54300 Human pan							
459	37	56.9	79	4	AAU40575	Aau40575 Propionib							
460	37	56.9	79	6	ABM37094	Abm37094 Propionib							
461	37	56.9	83	4	AAm94576	Aam94576 Human rep							
462	37	56.9	83	4	ABG60273	Abg60273 Human ova							

ALIGNMENTS

RESULT 1

ADR47260
ID ADR47260 standard; peptide; 10 AA.

XX ADR47260;

AC

XX 18-NOV-2004 (first entry)

DT

XX Human alpha2-antiplasmin cleaving enzyme internal sequence aa 210-219.

DE

XX antiinflammatory; cytostatic; vulnery; antiarteriosclerotic;

XX antithrombotic; vascular; cerebroprotective; pulmonary;

KW fibrin-related disorder; inflammatory condition; arthritis;

KW organ fibrosis; scarring; cancer; metastasis; atherothrombotic disease;

KW coronary artery thrombosis; stroke; pulmonary embolism; thrombosis;

KW human; alpha2-antiplasmin cleaving enzyme.

XX Homo sapiens.

OS

XX WO2004072240-A2.

FN

XX 26-AUG-2004.

PD

XX 07-FEB-2004; 2004WO-US003398.

PF

XX 07-FEB-2003; 2003US-0445774P.

XX 06-FEB-2004; 2004US-00774242.

PR

XX (MCKE/) MCKEE P A.

PA (LEEK/) LEE K N.

PA


```

XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 27; SEQ ID NO 30005; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridizes at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinemia, hyperlipidemia and hypercholesterolemia which is
XX CC associated with coronary heart disease. ABG47348-ABG5930 represent human
XX CC liver single exon encoded peptides of the invention. Note: the sequence
XX CC information for this patent does not appear in the printed specification
XX CC but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 51 AA;

Query Match 100.0%; Score 65; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSNPK 10
DB 7 YALWWSNPK 16

RESULT 7
ABG3294
ID ABG39294 standard; peptide; 51 AA.
XX AC ABG39294;
XX DT 19-AUG-2002 (first entry)
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 28959.
XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX KW primary ciliary dyskinesia; pulmonary hypertension;
XX KW hyaline membrane disease.
XX OS Homo sapiens.
XX OS
XX WO200186003-A2.

```

```

XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US000665.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX DR
XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX PT measure gene expression in human lung samples.
XX PS Claim 27; SEQ ID NO 28959; 634pp; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human lung comprising single exon nucleic acid probes having one of
XX CC 12614 nucleic acid sequences mentioned in the specification, or their
XX CC complements or the 12387 open reading frames derived from the 12614
XX CC probes. Also included are a microarray comprising the novel set of probes
XX CC ; the novel set of probes which hybridize at high stringency to a nucleic
XX CC acid expressed in the human lung; measuring gene expression in a sample
XX CC derived from human lung, comprising (a) contacting the array with a
XX CC collection of detectably labeled nucleic acids derived from human lung
XX CC mRNA, and (b) measuring the label detectably bound to each probe of the
XX CC array; identifying exons in a eukaryotic genome, comprising (a)
XX CC algorithmically predicting at least one exon from genomic sequences of
XX CC the eukaryote; and (b) detecting specific hybridisation of detectably
XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX CC having a fragment identical to the predicted exon, the probe is included
XX CC in the above mentioned microarray; assigning exons to a single gene,
XX CC comprising (a) identifying exons from genomic sequence by the method
XX CC above and (b) measuring the expression of each of the exons in several
XX CC tissues and/or cell types using hybridisation to a single exon
XX CC microarrays having a probe with the exon, where a common pattern of
XX CC expression of the exons in the tissues and/or cell types indicates that
XX CC the exons should be assigned to a single gene; a peptide comprising one
XX CC of 12011 sequences, mentioned in the specification, or encoded by the
XX CC probes/open reading frames (ORF). The probes are used for gene expression
XX CC analysis, and for identifying exons in a gene, particularly using human
XX CC lung derived mRNA and for the study of lung diseases such as asthma, lung
XX CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX CC Pudlak syndrome, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX CC present sequence is a peptide/protein encoded by a single exon probe of
XX CC the invention. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 51 AA;

Query Match 100.0%; Score 65; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSNPK 10
DB 7 YALWWSNPK 16

```


of A657D; and a composition comprising a FAP alpha dimer enzyme lacking amino acids 269-448 and comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, anti-parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from C, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of human soluble fibroblast activation protein (FAP) alpha dimer (corresponding to AEB94159 minus N-terminal amino acids 1-37).

Sequence 723 AA;

Query Match 100.0%; Score 65; DB 9; Length 723;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy .1 YALWWSPNGK 10

Db 173 YALWWSPNGK 182

RESULT 10

ID AEB94218 standard; protein; 734 AA.

AC AEB94218;

DT 06-OCT-2005 (first entry)

DE Human soluble (FAP) alpha dimer SEQ ID NO:61.

immune inhibition; fibroblast activation protein alpha dimer;
FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
immune disorder; neurological disease; autoimmune disease;
immunosuppressive; graft versus host disease; transplant rejection;
endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
musculoskeletal disease; allergy; antiasthmatic; asthma; antiasthmatic;
inflammation; respiratory disease; atherosclerosis; arteriosclerotic;
cardiovascular disease; metabolic disorder; hashimoto's disease;
antithyroid; endocrine disease; inflammatory bowel disease;
antifibrotic; gastrointestinal-gen.; gastrointestinal disease;
rheumatoid arthritis; antiinflammatory; multiple sclerosis; neuroprotective;
autoimmune hepatitis; antiinflammatory; hepatotropic;
systemic lupus erythematosus; dermatological; dermatological disease;
uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
hematological disease; rheumatic fever; antipyretic; Crohn's disease;
psoriasis; antipruritic; graves disease; antithyroid;
respiratory syncytial virus infection; respiratory-gen.; virucide.

OS Homo sapiens.

XX WO2005071073-A1.

PN 04-AUG-2005.

PD

PF 10-JAN-2005; 2005WO-US000709.

PR 09-JAN-2004; 2004US-0535577P.

XX (POIN-) POINT THERAPEUTICS INC.

XX Mclean PA, Jones B, Miller GT, Jesson MI;

DR WPI; 2005-564220/57.

Down-regulating an immune response comprises administering to a subject in need a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response.

Claim 134; SEQ ID NO 61; 177pp; English.

The invention relates to a method of down-regulating an immune response, which comprises administering to a subject a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response. Also included are the following: a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, where the composition is sterile and lacks an adjuvant; a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, and a non-adjuvant second agent; a composition comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising a FAP alpha dimer enzyme lacking amino acids 269-448 and comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, anti-parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from C, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of human soluble fibroblast activation protein (FAP) alpha dimer (corresponding to AEB94159 minus N-terminal amino acids 1-26).

Sequence 734 AA;

Query Match 100.0%; Score 65; DB 9; Length 734;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWWSPNGK 10

Db 184 YALWWSPNGK 193

RESULT 11

AEB94161
ID AEB94161 standard; protein; 750 AA.

XX AEB94161;

XX 06-OCT-2005 (first entry)

DE Human soluble fibroblast activation protein (FAP) alpha dimer.

CC	osteocarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c. autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of human soluble fibroblast activation protein (FAP) alpha dimer. Note: SEQ ID NO:3 AEB94160 only encodes amino acids 84 to 750 the present sequence.
XX	
SQ	Sequence 750 AA;
Query Match	100.0%; Score 65; DB 9; Length 750;
Best Local Similarity	100.0%; Pred. No. 0.12;
Matches 10; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 YALWWSPNGK 10
Db	199 YALWWSPNGK 208
RESULT 12	
AAW31963	
ID	AAW31963 standard; protein; 759 AA.
AC	AAW31963;
XX	
DT	21-JAN-1998 (first entry)
XX	
XX	Human fibroblast activation protein-alpha (FAP-alpha).
DE	
XX	Fibroblast activation protein-alpha; FAP-alpha; human;
KW	type II integral membrane proteins; epithelial cancer.
KW	
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	1..759
FT	/label= FAP-alpha
FT	/note= "Fibroblast activation protein alpha"
Modified-site	49..51
FT	/note= "Asn is N-glycosylated"
Region	191..209
FT	/label= fap-1
FT	/note= "Putative functional element; not defined further in specification"
Modified-site	214..216
FT	/note= "Asn is N-glycosylated"
Region	220..240
FT	/label= fap-2
FT	/note= "Putative functional element; not defined further in specification"
Region	510..521
FT	/label= fap-3
FT	/note= "Putative functional element; not defined further in specification"
Region	625..668
FT	/note= "alternative sequence found in AAW27438"
Modified-site	637..639
FT	/note= "Asn is N-glycosylated"
Modified-site	678..680
FT	/note= "Asn is N-glycosylated"
XX	
PX	WO9734927-A1.
XX	
PD	25-SEP-1997.
XX	
PF	12-MAR-1997; 97WO-US004215.
XX	
PR	18-MAR-1996; 96US-00619280.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.

(BOEH) BOEHRINGER INGELHEIM INT GMBH.

Zimmerman R, Park JS, Rettig W, Old LJ;
WPI: 1997-480164/44.
N-PSDB; AAT89286.

Dimeric fibroblast activation protein alpha - capable of degrading extracellular matrix protein, useful to treat epithelial cancers.

Example 5; Fig 1; 40pp; English.

This sequence represents the dimeric fibroblast activation protein alpha (FAP-alpha). The FAP-alpha molecule is capable of degrading extracellular matrix proteins and exhibits structural features typical of type II integral membrane proteins, including a large COOH-terminal extracellular domain, a hydrophobic transmembrane segment and a short cytoplasmic tail. The putative extracellular domain is described as containing five potential N-glycosylation sites, eleven conserved cysteine residues and three segments corresponding to highly conserved catalytic domains characteristic of serine proteases. Most common types of epithelial cancers including breast, lung, skin, pancreas and colorectal carcinomas contain abundant FAP-alpha reactive stromal fibroblasts which accompany tumour blood vessels. The induction of FAP-alpha fibroblasts at times and sites of tissue remodelling during foetal development, tissue repair and carcinogenesis is consistent with its putative role in normal fibroblast physiology. FAP alpha inhibitors, particularly a collagen derivative or (S)-valyl-pyrrolidine-2(R)-boronic acid, may be used to treat various epithelial cancers, e.g. breast, lung, skin, pancreas or colon cancer. Note: this sequence differs from the amino acid sequence of FAP-alpha found in AAW27438 between residues 625 to 670

Sequence 759 AA;

Query Match 100.0%; Score 65; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. NO. 0.12; Indels 0; Gaps 0
Matches 10; Conservative 0; Mismatches 0;

QY 1 YALWWSNGK 10
IIIIIIII
DB 210 YALWWSNGK 219

RESULT 13
AAW27438

ID AAW27438 standard; protein; 760 AA.

XX AC AAW27438;

XX XX

XX 21-JAN-1998 (first entry)

DE Human fibroblast activation protein-alpha (FAP-alpha).

XX Fibroblast activation protein-alpha; FAP-alpha; human;
XX type II integral membrane proteins; epithelial cancer.

XX OS Homo sapiens.

XX

FH Key Location/Qualifiers
FT Protein 1..760
FT /label= FAP-alpha
FT /note= "Fibroblast activation protein alpha"
FT Region 625..670
FT /note= "alternative sequence found in AAW31963"

XX MO9734927-Al.
XX 25-SEP-1997.

XX 12-MAR-1997; 97WO-US004215.
XX 18-MAR-1996; 96US-00619280.
XX

XX WPI; 2003-210381/20.
DR N-PSDB; ACC50144.
XX
XX Breast cancer diagnosis or treatment by comparing the level of expression
PT of a marker in a patient sample with that in the control non-breast
PT cancer sample.
XX
XX Claim 1; SEQ ID NO 136; 128pp; English.
XX
XX The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 760 AA;
SQ
Query Match 100.0%; Score 65; DB 6; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YALWWSNPK 10
Db 210 YALWWSNPK 219
|||||
RESULT 15
ADN95552
ID ADN95552 standard; protein; 760 AA.
XX
XX ADN95552;
XX
XX 01-JUL-2004 (first entry)
XX
XX Human BEC/LEC-related protein sequence SeqID475.
XX
XX growth; differentiation; blood endothelial cell; BEC;
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW lymphatic growth agent; VEGF-C; VEGF-D; angiogenic; cytostatic;
KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX
XX Homo sapiens.
OS
XX
XX WO2003080640-A1.
PN
XX
XX 02-OCT-2003.
PD
XX
XX 07-MAR-2003; 2003WO-US006900.
PF
XX
XX 07-MAR-2002; 2002US-0363019P.
PR
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX
XX (LICN) LICENTIA LTD.
PA
XX
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
PI
XX
XX WPI; 2003-876999/81.
DR
XX
XX N-PSDB; ADN95553.
DR
XX
XX Example 1; SEQ ID NO 475; 176pp; English.
PS
XX
XX This invention relates to a method of differentially modulating the
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic

CC endothelial cells (LEC) comprises contacting endothelial cells with a
CC composition comprising an agent that differentially modulates blood or
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC identifying a human subject with lymphoedema and with a mutation in at
CC least one allele of a gene encoding a LEC protein, where the mutation
CC correlates with lymphoedema in human subjects, and with the proviso that
CC the LEC protein is not VEGFR-3; and administering to the subject a
CC composition comprising a lymphatic growth agent selected from VEGF-C or
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC the development of compounds with an antiangiogenic, cytostatic, the method
CC vasotropic or antiinflammatory activity or for gene therapy. The method
CC is useful in modulating the growth or differentiation of blood
CC endothelial cells or lymphatic endothelial cells, in treating hereditary
CC lymphoedema, in screening for an endothelial cell disorder or
CC predisposition to the disorder or in monitoring the efficacy or toxicity
CC of a drug on endothelial cells. The agent is useful in manufacturing a
CC medicament for the differential modulation of blood vessel endothelial
CC cell or lymphatic vessel endothelial cell growth or differentiation. The
CC lymphatic growth agent may also be used in manufacturing a medicament for
CC the treatment of hereditary lymphoedema resulting from a mutation in a
CC LEC gene or of other diseases involving the lymphatic vessels, such as
CC various inflammatory diseases and cancer metastasis via the lymphatic
CC system. The present sequence is that of a human LEC/BEC differentially
CC expressed protein which is related to the method of the invention. Note:
CC This sequence does not appear in the specification but was obtained by
CC the indexer using the source data given in table 14 of the specification.
XX
XX Sequence 760 AA;
SQ
Query Match 100.0%; Score 65; DB 7; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YALWWSNPK 10
Db 210 YALWWSNPK 219
|||||
RESULT 16
ADQ21351
ID ADQ21351 standard; protein; 760 AA.
XX
XX ADQ21351;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human soft tissue sarcoma-upregulated protein - SEQ ID 4171.
DE
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
KW
XX
XX Homo sapiens.
OS
XX
XX WO2004048938-A2.
PN
XX
XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
PF
XX
XX 26-NOV-2002; 2002US-0429739P.
PR
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
DA
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX
XX WPI; 2004-441208/41.
DR
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 4171; 210pp; English.
PS
XX

CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytotatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC protein of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 760 AA;

Query Match 100.0%; Score 65; DB 8; Length 760;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10

DB 210 YALWWSPPNGK 219

RESULT 17

ID ADM14775 standard; protein; 760 AA.

XX AC ADM14775;

XX DT 07-APR-2005 (first entry)

XX DE Tumor-associated antigenic target TAT430.

XX DX Tumor-associated antigenic target; TAT430; cancer; neoplasm; cytostatic;
 KW ovary tumor; uterus tumor.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT FH Peptide 1..28

FT FT /label= Signal_peptide

FT FT Modified-site 10..15

FT FT /note= "N-myristoylation site"

FT FT Protein 29..760

FT FT /label= Mature_protein

FT FT Region 42..548

FT FT /note= "Dipeptidyl peptidase IV"

FT FT Modified-site 49..52

FT FT /note= "N-glycosylation site"

FT FT Modified-site 84..89

FT FT /note= "N-myristoylation site"

FT FT Modified-site 92..95

FT FT /note= "N-glycosylation site"

FT FT Modified-site 99..102

FT FT /note= "N-glycosylation site"

FT FT Modified-site 159..164

FT FT /note= "N-myristoylation site"

FT FT Modified-site 285..290

FT FT /note= "N-myristoylation site"

FT FT Modified-site 314..317

FT FT /note= "N-glycosylation site"

FT FT Modified-site 366..374

FT FT /note= "Tyrosine kinase phosphorylation site"

FT FT Modified-site 402..410

FT FT /note= "Tyrosine kinase phosphorylation site"

FT FT Modified-site 419..422

FT FT /note= "Amidation site"

FT FT Region 438..443

FT FT /note= "Cytochrome c family heme-binding site signature"

FT FT Modified-site 470..475

FT FT /note= "N-myristoylation site"

FT FT Modified-site 542..547

FT FT /note= "N-myristoylation site"

FT FT Region 552..629

FT FT /note= "Prolyl oligopeptidase family homology"

FT FT Modified-site 566..571

FT FT /note= "N-myristoylation site"

FT FT Active-site 599..629

FT FT /note= "Prolyl endopeptidase family serine active site"

FT FT Modified-site 626..631

FT FT /note= "N-myristoylation site"

FT FT Modified-site 636..639

FT FT /note= "Glycosaminoglycan attachment site"

FT FT Modified-site 639..644

FT FT /note= "N-myristoylation site"

FT FT Modified-site 679..682

FT FT /note= "N-glycosylation site"

FT FT Modified-site 738..743

FT FT /note= "N-myristoylation site"

PN WO2005003154-A2.

XX 13-JAN-2005.

XX 02-JUL-2004; 2004WO-US021353.

XX 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.

XX Ashkenazi A, Cairns B, Dowd P, Frantz G, French D, Gonzalez L;

XX Polakis P, Smith V, Wolf B, Wu TD, Zhang Z;

XX WPI; 2005-091764/10.

XX N-PSDB; ADM14768.

XX New Tumor-associated Antigenic Target polypeptides useful for diagnosing,

XX preventing or treating cell proliferative disorders such as cancer or

XX tumor.

XX Claim 1; SEQ ID NO 13; 141pp; English.

XX The present sequence is that of novel tumor-associated antigenic target

XX polypeptide TAT430. TAT polypeptides have been identified ADM14770-

XX ADM14776 whose expression is upregulated in particular tumor tissues.

XX Tissue expression profiling using GeneExpress indicated that TAT430 is

XX upregulated in breast, pancreatic, colon, adrenal, bladder, kidney,

XX liver, lung, lymphoid, ovarian, prostate and stomach tumor stroma as

XX compared to the corresponding healthy tissues. Upregulation in ovarian

XX tumor tissue was confirmed by quantitative analysis of TAT mRNA

XX expression, and upregulation in breast, pancreatic, colon, adrenal and

XX bladder tumors was confirmed by gene expression profiling in silico

XX (GEPIS) analysis. The TAT polypeptide may comprise the full-length

XX polypeptide, or may lack the associated signal peptide, or may comprise

XX an extracellular domain with or without its associated signal peptide, or

XX may be any of these fused to a heterologous polypeptide such as an

XX epitope tag, or the Cc region of an immunoglobulin. TAT polypeptides of

XX the invention, and their encoding nucleic acids, provide targets for the

XX diagnosis and therapy of cancer.

XX SQ Sequence 760 AA;

Query Match 100.0%; Score 65; DB 9; Length 760;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10

DB 210 YALWWSPPNGK 219

RESULT 18

AEB94159

ID AEB94159 standard; protein; 760 AA.

[illegible]

PS Disclosure; SEQ ID NO 6; 177pp; English.

XX The invention relates to a method of down-regulating an immune response,

CC which comprises administering to a subject a fibroblast activation

CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate

CC an immune response. Also included are the following: a composition

CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable

CC carrier, where the composition is sterile and lacks an adjuvant; a

CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically

CC acceptable carrier, and a non-adjuvant second agent; a composition

CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution

CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking

CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.

CC The method further comprises administering to the subject a second agent.

CC The second agent is an anti-inflammatory agent, immunosuppressant, or

CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-

CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild

CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation

CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The

CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a

CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid

CC substitution relative to wild type FAP alpha dimer. The amino acid

CC substitution is present in the beta-propeller domain, the catalytic

CC domain, or an N-linked glycosylation site and alters disulfide bond

CC formation. The immune response is an especially an IL-1 mediated

CC condition, abnormal immune response selected from inflammation,

CC autoimmune disease, sepsis, graft versus host disease, transplant

CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,

CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune

CC response is subsequent to an infection, such as an RSV infection. The

CC autoimmune disease is selected from c, autoimmune thyroiditis, systemic

CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,

CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,

CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple

CC sclerosis. The subject does not have cancer or a predisposition to

CC cancer. The present sequence represents the amino acid sequence of mouse

CC wild type fibroblast activation protein (FAP) alpha dimer.

XX

SQ Sequence 761 AA;

Query Match 92.3%; Score 60; DB 9; Length 761;

Best Local Similarity 90.0%; Pred. No. 0.74;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10

DB 210 YALWSPDGK 219

|||||:|

RESULT 20

AD117327

ID AD117327 standard; protein; 504 AA.

AC

XX AD117327;

XX

DT 15-APR-2004 (first entry)

XX

DE Polypeptide homologous to a human NOVX domain SeqID 863.

XX

KW NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes; inflammation;

KW autoimmune disorder; allergy; blood disorder;

KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

KW immunoglobulin (IgA) nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.

XX

OS Unidentified.

XX

PN WO200268649-A2.

XX

PD 06-SEP-2002.

XX

PF 31-JAN-2002; 2002WO-US002785.

XX

XX

PR 31-JAN-2001; 2001US-0265395P.

PR 31-JAN-2001; 2001US-0265412P.

PR 31-JAN-2001; 2001US-0265514P.

PR 02-FEB-2001; 2001US-0265517P.

PR 05-FEB-2001; 2001US-0266406P.

PR 07-FEB-2001; 2001US-0266767P.

PR 07-FEB-2001; 2001US-0266975P.

PR 08-FEB-2001; 2001US-0267057P.

PR 09-FEB-2001; 2001US-0267459P.

PR 15-FEB-2001; 2001US-0268233P.

PR 26-FEB-2001; 2001US-0268974P.

PR 27-FEB-2001; 2001US-0271664P.

PR 27-FEB-2001; 2001US-0271839P.

PR 02-MAR-2001; 2001US-0272788P.

PR 02-MAR-2001; 2001US-0273046P.

PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.

PR 14-MAR-2001; 2001US-0275950P.

PR 15-MAR-2001; 2001US-0275989P.

PR 15-MAR-2001; 2001US-0276448P.

PR 16-MAR-2001; 2001US-0276450P.

PR 16-MAR-2001; 2001US-0276397P.

PR 16-MAR-2001; 2001US-0276788P.

PR 20-MAR-2001; 2001US-0278652P.

PR 26-MAR-2001; 2001US-0278775P.

PR 26-MAR-2001; 2001US-0278778P.

PR 29-MAR-2001; 2001US-0279882P.

PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.

PR 11-APR-2001; 2001US-0282992P.

PR 11-APR-2001; 2001US-0283083P.

PR 23-APR-2001; 2001US-0285133P.

PR 23-APR-2001; 2001US-0285174P.

PR 03-MAY-2001; 2001US-0288327P.

PR 03-MAY-2001; 2001US-0288504P.

PR 29-MAY-2001; 2001US-0294047P.

PR 30-MAY-2001; 2001US-0294473P.

PR 08-JUN-2001; 2001US-0296964P.

PR 18-JUN-2001; 2001US-0298959P.

PR 19-JUN-2001; 2001US-0299324P.

PR 13-AUG-2001; 2001US-0312020P.

PR 16-AUG-2001; 2001US-0312889P.

PR 16-AUG-2001; 2001US-0312908P.

PR 21-AUG-2001; 2001US-0313390P.

PR 28-AUG-2001; 2001US-0315470P.

PR 31-AUG-2001; 2001US-0316447P.

PR 07-SEP-2001; 2001US-0318115P.

PR 07-SEP-2001; 2001US-0318118P.

PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.

PR 18-OCT-2001; 2001US-0330245P.

PR 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.

XX

PA (CURA-) CURAGEN CORP.

XX

XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;

PI Li L, Gangolli EA, Padigaru M, Anderson DW, Raetelli L, Miller CE;

PI Gerlach VL, Taupier RJ, Gusev VI, Colman SD, Wolenc AR, Pena CEA;

PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX

DR WPI; 2002-706998/76.

XX

XX New NOVX polypeptides and nucleic acids, useful for preventing or

PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,

PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX

XX Disclosure; SEQ ID NO 863; 1498pp; English.

PS

XX This invention relates to a novel nucleic acids, and encoded polypeptides

CC thereof, which have properties related to the stimulation of biochemical

or physiological responses in a cell, tissue, organ or organism.
Specifically, it refers to the use of biologically active fragments for
diagnostic and prognostic assays and furthermore in the treatment of
diverse pathological conditions. The present invention describes novel
human and murine NOVX proteins, as well as methods to modulate their
expression using antisense oligos, ribozymes and peptide nucleic acids.
The NOVX polypeptides, polynucleotides and antibodies are useful in
treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
atherosclerosis, cancer and diabetes. Furthermore, they may be used in
treating or preventing diseases such as inflammation, autoimmune
disorders, allergies, blood disorders, acquired immunodeficiency syndrome
(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
and epilepsy. Accordingly, these molecules have many activities including
cytostatic, anti-inflammatory, immunosuppressive, antiallergic,
antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,
relaxant and anticonvulsant. In addition, they are useful in screening
assays to identify small molecules that modulate or inhibit, for example,
neurogenesis, wound healing and angiogenesis. The nucleic acids are also
used as in chromosome mapping, tissue typing, preventive medicine and
pharmacogenomics. This polypeptide is a protein fragment that is
homologous to a domain of a human NOVX protein of the invention.

XX Sequence 504 AA;

Query Match 84.6%; Score 55; DB 5; Length 504;
Best Local Similarity 88.9%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWWSPNG 9
Db 171 YALWWSPDG 179
|||||:|

RESULT 21

ABP55609
ID ABP55609 standard; protein; 48 AA.

XX AC ABP55609;

XX DT 20-FEB-2003 (first entry)

XX DE Human DPP4 amino acid sequence #3.

XX KW DPP10; dipeptidyl peptidase; polyloligopeptidase; enzyme; asthma;
XX KW antiinflammatory; antiasthmatic; antipsoriatic; antiarthritic;
XX KW antirheumatic; vaccine; gene therapy; inflammatory disease;
XX KW inflammatory bowel disease; atopy; rheumatoid arthritis; psoriasis;
XX KW chromosome 2q14.

XX OS Homo sapiens.

XX PN WO200286113-A2.

XX PD 31-OCT-2002.

XX PF 24-APR-2002; 2002WO-GB001887.

XX PR 24-APR-2001; 2001GB-00010044.

XX PR 24-APR-2001; 2001GB-00010046.

XX PR 12-OCT-2001; 2001GB-00024575.

XX PR 12-OCT-2001; 2001GB-00024594.

XX PA (ISIS-) ISIS INNOVATIONS LTD.

XX PI Cookson WOOM, Moffat MF, Allen M, Lench N;

XX DR WPI; 2003-093132/08.

XX PT New nucleic acid sequence comprising DPP10 mRNA, useful for the

XX PT manufacture of a medicament for regulating DPP10 protein expression or

PT for preventing or treating inflammatory disease e.g., inflammatory bowel
PT disease.

XX PS Disclosure; Fig 17; 32pp; English.

XX CC The present invention describes a new isolated nucleic acid sequence (I)
CC comprising a DPP10 mRNA sequence. DPP10 is a dipeptidyl peptidase (also
CC known as polyloligopeptidase). (I) has antiinflammatory, antiasthmatic,
CC antipsoriatic, antiarthritic and antirheumatic activities, and can be
CC used in vaccines and gene therapy. A composition comprising (I) can be
CC used for the manufacture of a medicament for regulating DPP10 expression
CC or for preventing or treating inflammatory disease e.g., inflammatory
CC bowel disease, asthma, atopy, rheumatoid arthritis or psoriasis. (I) can
CC also be used in an assay for detecting or measuring DPP10 in a sample. A
CC host cell comprising (I) can be used for producing recombinant DPP10 gene
CC products, or in drug screening systems to identify agents for diagnosis
CC or treatment of individuals having or susceptible to inflammatory
CC disease. Human DPP10 is located on chromosome 2, more specifically
CC chromosome 2q14. ABQ84254 to ABQ84612 and ABP55569 to ABP55629 represent
CC sequences used in the exemplification of the present invention

XX SQ Sequence 48 AA;

Query Match 81.5%; Score 53; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPNG 9
Db 15 ALWWSPNG 22
|||||||

RESULT 22

AAR40916

ID AAR40916 standard; protein; 593 AA.

XX AC AAR40916;

XX DT 25-MAR-2003 (revised)

XX DT 05-FEB-1994 (first entry)

XX DE Sequence of a CD26 fragment lacking a portion of the carboxy terminal
XX DE region.

XX KW Human T cell activation antigen; monoclonal antibody Tai; CD26.

XX OS Synthetic.

XX PN WO9316102-A1.

XX PD 19-AUG-1993.

XX PF 09-APR-1992; 92WO-US002892.

XX PR 06-FEB-1992; 92US-00832211.

XX PA (DAND) DANA FARBBER CANCER INST INC.

XX PI Morimoto C, Schlossman SF, Tanaka T;

XX WI; 1993-272827/34.

XX PT Polypeptide fragments of CD26 - are capable of disrupting binding of CD45
XX PT and CD26 and thus interfering with T-cell activation.

XX PS Example; Page 46-48; 73pp; English.

XX CC C26 is a human T cell activation antigen originally identified by its
XX CC reactivity with the Mab Tai. C26 cDNA library was constructed from human
XX CC PHA-activated T cells using the CDM7 vector. Fragments of CD26 can be
XX CC prepd in the following manner. CD26 XbaI-SphI cDNA fragment is ligated to
XX CC the vector RCSR-alpha-26 XbaI-HindIII DNA fragment and the linker
XX CC AAQ46092. The linker introduces an in-frame stop codon that results in

CC the deletion of the segment of CD26 from AA 594 to the carboxy terminus
CC of the wild-type protein. This deletion mutant, shown in AAR40916, lacks
CC the putative catalytic site of CD26 and has a new carboxy terminus given
CC in AAR40917. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 593 AA;

Query Match 81.5%; Score 53; DB 2; Length 593;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ALWWSPPNG 9
| | | | |
DB 213 ALWWSPPNG 220

RESULT 23
AAR54614
ID AAR54614 standard; protein; 593 AA.
XX
AC AAR54614;
XX
DT 25-MAR-2003 (revised)
DT 09-DEC-1994 (first entry)
XX
DE Delta594-C CD26.
XX Human; T cell activation antigen; CD26; analogues; deletion; soluble;
KW signal peptidase; immune-stimulating; response-stimulating; AIDS;
KW immunosuppression; AIDS-related complex.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 593 /note= "Position of delta594-C deletion"
FT
PN WO9409132-A1.
XX
PD 28-APR-1994.
XX
PF 19-AUG-1993; 93WO-US007923.
XX
PR 21-AUG-1992; 92US-00934162.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Morimoto C, Schlossman S, Tanaka T;
XX WPI; 1994-151317/18.
XX
PT Polypeptide fragments and analogues of CD26 and encoding nucleic acid -
PT useful for stimulating immune response, e.g. for treatment of AIDS to
PT counteract immunosuppressive drug, and as vaccine adjuvant.
XX
PS Disclosure; Page 56-58; 85pp; English.
XX

CC The sequences given in AAR54612-14 represents analogues of the human T
CC cell activation antigen CD26 which have internal deletions. The analogues
CC pref. lack residues 3-9 or 24-34. These analogues are soluble under
CC physiological conditions and lack enough amino acid residues to render
CC them susceptible to cleavage by signal peptidase. The peptide fragments
CC and analogues are useful as immune or response- stimulating therapeutics,
CC eg. they may be used for treatment of disease conditions characterised by
CC immunosuppression, eg. AIDS or AIDS-related complex, other virally or
CC environmentally-induced conditions, and certain congenital immune
CC deficiencies. The peptides can be employed to increase immune function
CC which has been impaired by use of immunosuppressive drugs, such as certain
CC chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 593 AA;

Query Match 81.5%; Score 53; DB 2; Length 593;

Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ALWWSPPNG 9
| | | | |
DB 213 ALWWSPPNG 220

RESULT 24
ADO71642
ID ADO71642 standard; protein; 688 AA.
XX
AC ADO71642;
XX
DT 26-AUG-2004 (first entry)
XX

DE Amino acid sequence of a human CD26 protein.

XX CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
KW dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
XX

OS Homo sapiens.

XX WO2004045497-A2.

XX 03-JUN-2004.

XX 15-MAY-2003; 2003WO-US015499.

XX 17-MAY-2002; 2002US-0381606P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Dang NH, Morimoto C;

XX WPI; 2004-420511/39.

XX N-PSDB; ADO71641.

XX Use of a CD26 composition, and a chemotherapeutic and/or a
PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
PT regression or tumor necrosis.

XX Claim 23; Page 170-171; 182pp; English.

XX The specification describes a CD26 composition which, in conjunction with
CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
CC is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
CC topoisomerase II inhibitor. The CD26 composition of the invention is
CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
CC a cell, killing a cancer cell, potentiating the effect of a
CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
CC inducing tumour regression or tumour necrosis. The CD26 composition is
CC further useful for increasing topoisomerase II expression in a cell, for
CC activating an antigen-presenting cell, or for potentiating immune
CC responses of an animal. The present sequence represents a CD26 protein,
CC and is encoded by vectors which are used to produce compositions of the
CC invention.

XX Sequence 688 AA;

Query Match 81.5%; Score 53; DB 8; Length 688;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
| | | | |

DB 213 ALWWSPPNG 220

RESULT 25
ADO40240
XX ADO40240 standard; protein; 736 AA.
XX
XX ADO40240;
XX
XX 12-AUG-2004 (first entry)
XX Human DPP-IV extracellular domain protein SEQ ID NO:2.
XX
XX crystal; mammalian dipeptidyl-peptidase IV extracellular domain;
XX dipeptidyl-peptidase IV extracellular domain;
XX DPP-IV extracellular domain; three-dimensional structure; antidiabetic;
XX anorectic; cytostatic; type I diabetes; type II diabetes; IGT; obesity;
XX cancer; human; DPP-IV; enzyme; protein co-ordinate data; EC 3.4.14.5.
XX
XX Homo sapiens.
XX
XX EPI422293-A1.
XX
XX 26-MAY-2004.
XX
XX 17-NOV-2003; 2003EP-00026169.
XX
XX 25-NOV-2002; 2002EP-00026367.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX Hennig M, Loeffler BM, Thoma R;
XX
XX WPI; 2004-413363/39.
XX
XX N-PSDB; ADO40239.
XX
XX New crystal of an extracellular domain of mammalian dipeptidyl-peptidase
XX IV (DPP-IV) useful for identifying or designing inhibitors of DPP-IV
XX activity.
XX
XX Claim 31; SEQ ID NO 2; 215pp; English.
XX
XX The present invention describes a crystal (I) of the extracellular domain
XX of mammalian dipeptidyl-peptidase (DPP)-IV (EC 3.4.14.5). Also described:
XX (1) a co-crystal of the extracellular domain of mammalian DPP-IV and a
XX ligand bound to its active site; (2) a co-crystal of the extracellular
XX domain of mammalian DPP-IV and a ligand bound to an allosteric binding
XX site; (3) a co-crystal of the extracellular domain of mammalian DPP-IV
XX and HgCl₂; (4) crystallising (M1) mammalian DPP-IV; (5) co-crystallising
XX (M2) mammalian DPP-IV and an active site ligand; (6) a crystal produced
XX by (M1) and (M2); (7) determining the three-dimensional structure of a
XX crystallised extracellular domain of mammalian DPP-IV to a resolution of
XX 3.5-2.1 angstrom or better; (8) a machine-readable data storage medium
XX comprising a data storage material encoded with machine readable data
XX which, when using a machine programmed with instructions for using the
XX data, displays a graphical three-dimensional representation of a molecule
XX or molecular complex comprising at least a portion of the extracellular
XX domain of mammalian DPP-IV comprising a fully defined sequence (SEQ ID
XX NO:2, S1) of 736 amino acids, where the extracellular domain comprising
XX the ligand binding active site being defined by a set of points having a
XX root mean square deviation of less than about 1.5 angstrom from points
XX representing the backbone atoms of the amino acids as represented by
XX structure coordinates as given in the specification; (9) a compound (II)
XX identified by using (I); (10) a pharmaceutical composition (III)
XX comprising (I) and a carrier; (11) an isolated nucleic acid sequence (IV)
XX encoding the soluble extracellular domain of DPP-IV comprising a fully
XX defined sequence (SEQ ID NO:1, S2) of 221 nucleotides; (12) a nucleic
XX acid construct (V) comprising an expression vector and (IV); (13) a host
XX cell (VI) transformed with (V); (14) producing the soluble extracellular
XX domain of DPP-IV, involves culturing (VI) under conditions permitting the
XX expression of the soluble extracellular domain of DPP-IV by (VI); and
XX (15) a polypeptide comprising the soluble extracellular domain of (S1).
XX DPP-IV has antidiabetic, anorectic and cytostatic activities. (I) is
XX useful for identifying a compound that interacts with DPP-IV. The
XX compound interacts with the active site of DPP-IV. The compound interacts

CC with an allosteric binding site of DPP-IV. The compound is an inhibitor
CC of DPP-IV activity. (I) is useful for the identification and/or design of
CC inhibitors of DPP-IV activity. (II) is useful as a therapeutic active
CC substance, in particular for the treatment of diabetes type I, diabetes
CC type II, IGT, obesity and cancer. (III) is useful for the manufacture of a
CC medicament for the treatment of above mentioned disease. The present
CC sequence represents the extracellular domain of human DPP-IV, which is
CC used in the exemplification of the present invention.
XX
XX

SQ Sequence 736 AA;

Query Match 81.5%; Score 53; DB 8; Length 736;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALWWSPPNG 9
| | | | | | | |
Db 183 ALWWSPPNG 190

RESULT 26

AAR54613

ID AAR54613 standard; protein; 739 AA.

XX AAR54613;

XX 25-MAR-2003 (revised)

DT 09-DEC-1994 (first entry)

XX Delta24-34 CD26.

XX Human; T cell activation antigen; CD26; analogues; deletion; soluble;
XX signal peptidase; immune-stimulating; response-stimulating; AIDS;
XX immunosuppression; AIDS-related complex.

XX Homo sapiens.

XX Location/Qualifiers

PH Key

FT Misc-difference 23..24 /note= "Position of delta24-34 deletion"

XX WO9409132-A1.

XX 28-APR-1994.

XX 19-AUG-1993; 93WO-US007923.

XX 21-AUG-1992; 92US-00934162.

XX (DAND) DANA FARBER CANCER INST INC.

XX Morimoto C, Schlossman S, Tanaka T;

XX WPI; 1994-151317/18.

XX Polypeptide fragments and analogues of CD26 and encoding nucleic acid -
XX useful for stimulating immune response, e.g. for treatment of AIDS to
XX counteract immunosuppressive drug, and as vaccine adjuvant.

XX Claim 4; Page 52-54; 85pp; English.

XX The sequences given in AAR54612-14 represents analogues of the human T
XX cell activation antigen CD26 which have internal deletions. The analogues
XX pref. lack residues 3-9 or 24-34. These analogues are soluble under
XX physiological conditions and lack enough amino acid residues to render
XX them susceptible to cleavage by signal peptidase. The peptide fragments
XX and analogues are useful as immune or response- stimulating therapeutics,
XX eg. they may be used for treatment of disease conditions characterised by
XX immunosuppression, eg. AIDS or AIDS-related complex, other virally or
XX environmentally-induced conditions, and certain congenital immune
XX deficiencies. The peptides can be employed to increase immune function
XX which has been impaired by use of immunosuppressive drugs, such as certain
XX chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)

```
XX SQ Sequence 739 AA;
Query Match 81.5%; Score 53; DB 2; Length 739;
Best Local Similarity 100.0%; Pred. No. 8.9; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 2 ALWSPNG 9
|||||
DB 202 ALWSPNG 209

RESULT 27
AAR54612
ID AAR54612 standard; protein; 759 AA.
XX AC AAR54612;
XX 25-MAR-2003 (revised)
DT 09-DEC-1994 (first entry)
DE Delta3-9 CD26.
XX Human; T cell activation antigen; CD26; analogues; deletion; soluble;
KW signal peptidase; immune-stimulating; response-stimulating; AIDS;
KW immunosuppression; AIDS-related complex.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Misc-difference 2..3 /note= "Position of delta3-9 deletion"
FT FT
XX WO9409132-A1.
XX 28-APR-1994.
XX 19-AUG-1993; 93WO-US007923.
XX 21-AUG-1992; 92US-00934162.
XX (DAND ) DANA FARBER CANCER INST INC.
XX Morimoto C, Schlossman S, Tanaka T;
XX WPI; 1994-151317/18.
XX Polypeptide fragments and analogues of CD26 and encoding nucleic acid -
PT useful for stimulating immune response, e.g. for treatment of AIDS to
PT counteract immunosuppressive drug, and as vaccine adjuvant.
XX Claim 3; Page 49-52; 85pp; English.
XX The sequences given in AAR54612-14 represents analogues of the human T
CC cell activation antigen CD26 which have internal deletions. The analogues
CC pref. lack residues 3-9 or 24-34. These analogues are soluble under
CC physiological conditions and lack enough amino acid residues to render
CC them susceptible to cleavage by signal peptidase. The peptide fragments
CC and analogues are useful as immune or response- stimulating therapeutics,
CC eg. they may be used for treatment of disease conditions characterised by
CC immunosuppression, eg. AIDS or AIDS-related complex, other virally or
CC environmentally-induced conditions, and certain congenital immune
CC deficiencies. The peptides can be employed to increase immune function
CC which has been impaired by use of immunosuppressive drugs, such as certain
CC chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 759 AA;
Query Match 81.5%; Score 53; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ALWSPNG 9

|||||
DB 206 ALWSPNG 213

RESULT 28
AAR40909
ID AAR40909 standard; protein; 766 AA.
XX AC AAR40909;
XX 25-MAR-2003 (revised)
DT 05-FEB-1994 (first entry)
DE Sequence encoded by human CD26 cDNA.
XX Human T cell activation antigen; monoclonal antibody Tai.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Region 7..28 /label= hydrophobic
FT Region 29..323 /label= N-terminal glycosylated region of extracellular
FT domain /note= "8 sites for N-linked glycans"
FT Region 324..551 /label= Cysteine rich region of extracellular domain
FT /note= "1 N-linked glycosylation site"
FT Region 552..766 /label= C-terminal region of extracellular domain
FT /note= "1 N-linked glycosylation site & 1 catalytic site"
FT Active-site 627..631 /label= active site of serine protease/esterase
FT /note= "fits the consensus sequence GXSGXG"
XX WO9316102-A1.
XX 19-AUG-1993.
XX 09-APR-1992; 92WO-US002892.
XX 06-FEB-1992; 92US-00832211.
XX (DAND ) DANA FARBER CANCER INST INC.
XX Morimoto C, Schlossman SF, Tanaka T;
XX WPI; 1993-272827/34.
XX N-PSDB; AAQ46089.
XX Polypeptide fragments of CD26 - are capable of disrupting binding of CD45
PT and CD26 and thus interfering with T-cell activation.
XX Disclosure; Page 39-43; 73pp; English.
XX C26 is a human T cell activation antigen originally identified by its
CC reactivity with the MAb Tai. C26 cDNA library was constructed from human
CC PHA-activated T cells using the CDW7vector. The hydrophobic N-terminal of
CC the predicted CD26 polypeptide has the characteristics of a signal
CC sequence of the type II membrane protein, which is reinforced by the
CC observation that potential N-glycosylation sites are located in the
CC carboxy side of the hydrophobic core. Therefore the N-terminal 6 AAs are
CC predicted to be cytoplasmic, the next 22 AAs are predicted to transverse
CC the cytoplasmic membrane, and the 738 C-terminal AAs constitute the
CC predicted extracellular domain. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX SQ Sequence 766 AA;
Query Match 81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


PR 31-OCT-2000; 2000US-0244524P.
XX (UYVA-) UNIV VANDERBILT.
XX Brown NJ;
PI WPI; 2002-627422/67.
XX N-PSDB; AAL44351.
DR
XX
PT Diagnosing susceptibility to developing an angiotensin converting enzyme
PT (ACE) inhibitor- or a vasopeptidase inhibitor-associated angioedema, by
PT measuring levels of dipeptidyl peptidase IV or aminopeptidase P enzyme
PT activities.
XX
PS Disclosure; Page 62-64; 71pp; English.
XX
CC The invention comprises a method of identifying a subject that is
CC susceptible to developing angioedemic conditions during a course of
CC treatment. The method of the invention involves administering an
CC angiotensin converting enzyme (ACE) inhibitor or a vasopeptidase
CC inhibitor and determining dipeptidyl peptidase IV (DPP IV) enzyme
CC activity or aminopeptidase P (APP) enzyme activity. The method of the
CC invention is useful for identifying/diagnosing susceptibility to
CC developing angioedemic conditions during a course of treatment involving
CC the administration of an ACE inhibitor or a vasopeptidase inhibitor. The
CC method is particularly useful during the treatment of a subject that is
CC in need of or taking an ACE inhibitor and/or a vasopeptidase inhibitor,
CC which are commonly used in the treatment of hypertension, diabetes,
CC cardiac disease and renal diseases. The present amino acid sequence
CC represents the human dipeptidyl peptidase IV (DPP IV)
XX
SQ Sequence 766 AA;
Query Match 81.5%; Score 53; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 ALWWSPPNG 9
DB 213 ALWWSPPNG 220
RESULT 34
ABP55629
ID ABP55629 standard; protein; 766 AA.
XX
AC ABP55629;
XX
DT 20-FEB-2003 (first entry)
XX
XX Human dpp4 protein sequence.
XX
XX DPP10; dipeptidyl peptidase; prolololigopeptidase; enzyme; asthma;
KW antiinflammatory; antiasthmatic; antipruritic; antiasthmatic; infection;
KW antirheumatic; vaccine; gene therapy; inflammatory disease;
KW inflammatory bowel disease; atopy; rheumatoid arthritis; psoriasis;
KW chromosome 2q14.
XX
XX Homo sapiens.
OS
XX
PN WO200286113-A2.
XX
PD 31-OCT-2002.
XX
XX 24-APR-2002; 2002WO-GB001887.
PF
XX 24-APR-2001; 2001GB-00010044.
PR 24-APR-2001; 2001GB-00010046.
PR 12-OCT-2001; 2001GB-00024575.
PR 12-OCT-2001; 2001GB-00024594.
XX
XX (ISIS-) ISIS INNOVATIONS LTD.
PA

PI Cookson WOCM, Moffat MF, Allen M, Lench N;
XX WPI; 2003-093132/08.
XX
PT New nucleic acid sequence comprising DPP10 mRNA, useful for the
PT manufacture of a medicament for regulating DPP10 protein expression or
PT for preventing or treating inflammatory disease e.g., inflammatory bowel
PT disease.
XX
PS Example 2; Fig 23; 321pp; English.
XX
CC The present invention describes a new isolated nucleic acid sequence (I)
CC comprising a DPP10 mRNA sequence. DPP10 is a dipeptidyl peptidase (also
CC known as prolololigopeptidase). (I) has antiinflammatory, antiasthmatic,
CC antipruritic, antiasthmatic and antirheumatic activities, and can be
CC used in vaccines and gene therapy. A composition comprising (I) can be
CC used for the manufacture of a medicament for regulating DPP10 expression
CC or for preventing or treating inflammatory disease e.g., inflammatory
CC bowel disease, asthma, atopy, rheumatoid arthritis or psoriasis. (I) can
CC also be used in an assay for detecting or measuring DPP10 in a sample. A
CC host cell comprising (I) can be used for producing recombinant DPP10 gene
CC products, or in drug screening systems to identify agents for diagnosis
CC or treatment of individuals having or susceptible to inflammatory
CC disease. Human DPP10 is located on chromosome 2, more specifically
CC chromosome 2q14. ABQ84254 to ABQ84612 and ABP55569 to ABP55629 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 766 AA;
Query Match 81.5%; Score 53; DB 6; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 ALWWSPPNG 9
DB 213 ALWWSPPNG 220
RESULT 35
ABP56700
ID ABP56700 standard; protein; 766 AA.
XX
AC ABP56700;
XX
DT 27-MAR-2003 (first entry)
XX
XX Human liver dipeptidyl peptidase IV (DPP4) protein SEQ ID NO:7.
XX
XX Tissue-specific endothelial membrane protein; antibacterial; virucide;
KW fungicide; antiparasitic; anticonvulsant; neuroleptic; cytostatic;
KW antiparkinsonian; nootropic; neuroprotective; antiasthmatic; infection;
KW antidiabetic; hypotensive; nephrotoxic; cancer; Parkinson's disease;
KW gene therapy; epilepsy; schizophrenia; diabetes; hypertension; arthritis;
KW Alzheimer's disease; asthma; diabetes; hypertension; arthritis;
KW polycystic kidney disease; tissue-specific luminal expressed protein;
KW inflammatory bowel disease; enzyme.
XX
XX Homo sapiens.
OS
XX
PN WO2002100336-A2.
XX
PD 19-DEC-2002.
XX
XX 07-JUN-2002; 2002WO-US018185.
PF
XX 08-JUN-2001; 2001US-0297021P.
PR 12-JUL-2001; 2001US-0305117P.
XX
XX (TARG-) TARGET PROTEIN TECHNOLOGIES INC.
PA
XX Roben PW, Stevens AC;
PI
XX WPI; 2003-167367/16.
XX

DR N-PSDB; ABZ22557.

XX Delivering a therapeutic agent (e.g. immunosuppressant) to a specific

PT tissue, comprises administering an amount of a therapeutic complex

PT comprising a therapeutic group, a ligand that binds to a tissue-specific

PT protein, and a linker.

XX Example 2; Page 106-107; 123pp; English.

XX The present invention describes a method (M1) for delivering a

CC therapeutic agent to a specific tissue, comprising administering an

CC amount of a therapeutic complex comprising a ligand which binds to a

CC tissue-specific luminally expressed protein, a therapeutic group, and a

CC linker which links the group to the ligand. Tissue-specific endothelial

CC membrane proteins from the present invention can have antibacterial,

CC virucide, fungicide, antiparasitic, anticonvulsant, neuroleptic,

CC cytostatic, antiparkinsonian, nootropic, neuroprotective, antiasthmatic,

CC antidabetic, hypotensive, nephrotropic, antiarthritic and

CC antiinflammatory activities, and can be used in gene therapy. (M1) is

CC useful in targeting pharmaceuticals or other therapeutics to specific

CC tissues using tissue-specific endothelial membrane proteins. A

CC therapeutic complex may be used to treat or diagnose any disease for

CC which a tissue- or organ-specific treatment would be efficacious, such as

CC in cases of infections (e.g. bacterial, viral, fungal and parasitic),

CC epilepsy, schizophrenia, cancer, Parkinson's disease, Alzheimer's

CC disease, asthma, diabetes, hypertension, polycystic kidney disease,

CC arthritis, and inflammatory bowel disease. The present sequence

CC represents a human liver dipeptidyl peptidase IV (DPP4), which is used in

CC an example from the present invention

XX Sequence 766 AA;

XX Query Match 81.5%; Score 53; DB 6; Length 766;

XX Best Local Similarity 100.0%; Pred. No. 9.2; Mismatches 0; Gaps 0;

XX Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 2 ALWNSPNG 9

DB 213 ALWNSPNG 220

RESULT 36

ADD14045

ID ADD14045 standard; protein; 766 AA.

XX ADD14045;

AC ADD14045;

XX 01-JAN-2004 (first entry)

XX Human arc biomarker polypeptide SEQ ID NO:234.

XX predictor set; protein tyrosine kinase activity modulator;

XX protein tyrosine kinase pathway; protein tyrosine kinase; cytosolic;

XX gene therapy; drug sensitivity; genetic profile; cancer; human.

XX Homo sapiens.

XX WO2003062395-A2.

XX 31-JUL-2003.

XX 17-JAN-2003; 2003WO-US001981.

XX 18-JAN-2002; 2002US-0350061P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Huang F, Fairchild CR, Lee FY, Shaw P;

XX WPI; 2003-636735/60.

XX N-PSDB; ADD14640.

XX New polynucleotides and polypeptides for predicting the activity of

PT compounds that interact with protein tyrosine kinases and/or protein

PT tyrosine kinase pathways.

XX Claim 10; SEQ ID NO 234; 139pp; English.

XX The present invention describes a predictor set comprising a plurality of

CC polynucleotides or polypeptides whose expression pattern is predictive of

CC the response of cells to treatment with a compound that modulates protein

CC tyrosine kinase activity or members of the protein tyrosine kinase

CC pathway. Also described: (1) predicting whether a compound is capable of

CC modulating the activity of cells, comprising obtaining a sample of cells,

CC determining whether the cells express a plurality of markers, and

CC correlating the expression of the markers to the compound's ability to

CC modulate the activity of the cells; (2) a plurality of cell lines for

CC identifying polynucleotides and polypeptides whose expression levels

CC correlate with compound sensitivity or resistance of cells associated

CC with a disease state; and (3) identifying polynucleotides and

CC polypeptides that predict compound sensitivity or resistance of cells

CC associated with a disease state, comprising subjecting the plurality of

CC cell lines to one or more compounds, analysing the expression pattern of

CC a microarray of polynucleotides or polypeptides, and selecting

CC polynucleotides or polypeptides that predict the sensitivity or

CC resistance of cells associated with a disease state by using the

CC expression pattern of the microarray. The polynucleotides and

CC polypeptides have cytostatic activities, and can be used in gene therapy.

CC The polynucleotides and polypeptides are useful in predicting the

CC activity of compounds that interact with protein tyrosine kinases and/or

CC protein tyrosine kinase pathways. These may be used in determining drug

CC sensitivity in patients to allow the development of individualized

CC genetic profiles which aid in treating diseases and disorders (e.g.

CC cancer) based on patient response at a molecular level. The present

CC sequence is used in the exemplification of the present invention.

XX Sequence 766 AA;

XX Query Match 81.5%; Score 53; DB 7; Length 766;

XX Best Local Similarity 100.0%; Pred. No. 9.2; Mismatches 0; Gaps 0;

XX Matches 8; Conservative 0; Indels 0; Gaps 0;

QY 2 ALWNSPNG 9

DB 213 ALWNSPNG 220

RESULT 37

ADD27855

ID ADD27855 standard; protein; 766 AA.

XX ADD27855;

AC ADD27855;

XX 15-JAN-2004 (first entry)

XX Human dipeptidyl peptidase IV (DPPIV).

XX Mucosal inflammation; rhinitis; sinusitis; exopeptidase; substance P; SP;

XX neurokinin 1 receptor; NK1 receptor; allergy; asthma; antiallergic;

XX antiinflammatory; antiasthmatic; human; dipeptidyl peptidase IV; DPPIV;

XX enzyme.

XX Homo sapiens.

XX US2003165489-A1.

XX 04-SEP-2003.

XX 27-NOV-2001; 2001US-00993959.

XX 28-FEB-2001; 2001US-00794236.

XX (BMRA-) BMRA CORP BV.

XX Grouzmann E, Lacroix J, Monod M;

XX

DR WPI; 2003-811386/76.
XX Treatment of patient for mucosal inflammation associated with rhinitis
PT and/or sinusitis involves intranasally administering peptidase that
PT cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to
PT neurokinin 1 receptor.
XX Disclosure; SEQ ID NO 1; 14pp; English.
XX The present invention relates to a method of treating a patient for
CC mucosal inflammation associated with rhinitis and/or sinusitis. The
CC method comprises intranasally administering to the patient a peptidase
CC that cleaves at Xaa-Pro sequences or an agent that inhibits the binding
CC of substance P (SP) to the neurokinin 1 (NK1) receptor. The peptidase is
CC an exopeptidase, preferably selected from human dipeptidyl peptidase IV
CC (DPPIV), human quiescent cell proline dipeptidase, human dipeptidyl
CC peptidase 8, or human attractin. The method is useful for treating a
CC patient for mucosal inflammation associated with rhinitis and/or
CC sinusitis which are the result of allergies or asthma. The invention
CC provides an effective treatment of the inflammation associated with both
CC rhinitis and sinusitis. The present sequence represents human DPPIV.
XX
SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 7; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWNSPNG 9
DB 213 ALWNSPNG 220

RESULT 38
ADD46934
ID ADD46934 standard; protein; 766 AA.
XX AC ADD46934;
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX Human Protein AAA52308, SEQ ID NO 12620.
DE
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
OS Unidentified.
XX
XX WO2003016475-A2.
PN
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; AAA52308.
DR
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
PS

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 7; Length 766;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWNSPNG 9
DB 213 ALWNSPNG 220

RESULT 39
ADN39272
ID ADN39272 standard; protein; 766 AA.
XX AC ADN39272;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:590.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX
XX Homo sapiens.
OS
XX WO2003042661-A2.
PN
XX
PD 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-032464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR

OS	Homo sapiens.
XX	WO2003042661-A2.
XX	22-MAY-2003.
XX	13-NOV-2002; 2002WO-US036810.
XX	13-NOV-2001; 2001US-0350666P.
PR	21-NOV-2001; 2001US-0332464P.
PR	29-NOV-2001; 2001US-0334393P.
PR	03-DEC-2001; 2001US-0335394P.
PR	14-DEC-2001; 2001US-0340376P.
PR	08-JAN-2002; 2002US-0347211P.
PR	10-JAN-2002; 2002US-0347349P.
PR	08-FEB-2002; 2002US-0355250P.
PR	13-FEB-2002; 2002US-0356714P.
PR	20-FEB-2002; 2002US-0359077P.
PR	20-MAR-2002; 2002US-0368809P.
PR	04-APR-2002; 2002US-0370110P.
PR	12-APR-2002; 2002US-0372246P.
PR	05-JUN-2002; 2002US-0386614P.
PR	16-JUL-2002; 2002US-0396839P.
PR	22-JUL-2002; 2002US-0397775P.
PR	22-JUL-2002; 2002US-0397845P.
PR	09-SEP-2002; 2002US-0409450P.
XX	(EOSB-) EOS BIOTECHNOLOGY INC.
PA	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX	WPI; 2003-468649/44.
DR	N-PSDB; ADN39603.
XX	Determining the presence or absence of a pathological cell in a patient,
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting
PT	a nucleic acid in a biological sample.
XX	Claim 12; SEQ ID NO A204; 1385pp; English.
XX	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC	whose expression is upregulated or downregulated in specific cancers or
CC	other diseases such as angiogenic or fibrotic disorders, and to methods
CC	of determining the presence or absence of a pathological cell in a
CC	patient by detecting a nucleic acid at least 80% identical to those of
CC	the invention or by detecting a polypeptide of the invention. The
CC	invention also relates to expression vectors and host cells comprising a
CC	nucleic acid of the invention; antibodies which specifically bind a
CC	polypeptide of the invention; use of such antibodies for drug targeting;
CC	and methods of screening for modulators of activity or expression of the
CC	antibodies and nucleic acids. The nucleic acids, polypeptides,
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC	neovascularisation syndromes, scarring and uterine fibroids. They may
CC	also be useful in wound healing and in contraception. The present
CC	sequence represents a polypeptide of the invention.
XX	Sequence 766 AA;
Query Match	81.5%; Score 53; DB 7; Length 766;
Best Local Similarity	100.0%; Pred. No. 9.2;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2 ALWWSPNG 9
Db	
	213 ALWWSPNG 220
RESULT 40	
ID	ADN39604
ID	ADN39604 standard; protein; 766 AA.
XX	AC ADN39604;
XX	17-JUN-2004 (first entry)
DE	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A204.
XX	Human; differential expression; cancer; angiogenic disorder;
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW	inflammatory disease; autoimmune disease;
KW	retinal neovascularisation syndrome; scarring; uterine fibroid;
KW	detection; diagnosis; prognosis; drug screening; drug targeting;
KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW	vulnerable; gene therapy; vaccine.
XX	

OS	Homo sapiens.
XX	WO2003042661-A2.
XX	22-MAY-2003.
XX	13-NOV-2002; 2002WO-US036810.
XX	13-NOV-2001; 2001US-0350666P.
PR	21-NOV-2001; 2001US-0332464P.
PR	29-NOV-2001; 2001US-0334393P.
PR	03-DEC-2001; 2001US-0335394P.
PR	14-DEC-2001; 2001US-0340376P.
PR	08-JAN-2002; 2002US-0347211P.
PR	10-JAN-2002; 2002US-0347349P.
PR	08-FEB-2002; 2002US-0355250P.
PR	13-FEB-2002; 2002US-0356714P.
PR	20-FEB-2002; 2002US-0359077P.
PR	20-MAR-2002; 2002US-0368809P.
PR	04-APR-2002; 2002US-0370110P.
PR	12-APR-2002; 2002US-0372246P.
PR	05-JUN-2002; 2002US-0386614P.
PR	16-JUL-2002; 2002US-0396839P.
PR	22-JUL-2002; 2002US-0397775P.
PR	22-JUL-2002; 2002US-0397845P.
PR	09-SEP-2002; 2002US-0409450P.
XX	(EOSB-) EOS BIOTECHNOLOGY INC.
PA	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX	WPI; 2003-468649/44.
DR	N-PSDB; ADN39603.
XX	Determining the presence or absence of a pathological cell in a patient,
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting
PT	a nucleic acid in a biological sample.
XX	Claim 12; SEQ ID NO 590; 1385pp; English.
XX	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC	whose expression is upregulated or downregulated in specific cancers or
CC	other diseases such as angiogenic or fibrotic disorders, and to methods
CC	of determining the presence or absence of a pathological cell in a
CC	patient by detecting a nucleic acid at least 80% identical to those of
CC	the invention or by detecting a polypeptide of the invention. The
CC	invention also relates to expression vectors and host cells comprising a
CC	nucleic acid of the invention; antibodies which specifically bind a
CC	polypeptide of the invention; use of such antibodies for drug targeting;
CC	and methods of screening for modulators of activity or expression of the
CC	antibodies and nucleic acids. The nucleic acids, polypeptides,
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC	neovascularisation syndromes, scarring and uterine fibroids. They may
CC	also be useful in wound healing and in contraception. The present
CC	sequence represents a polypeptide of the invention.
XX	Sequence 766 AA;
Query Match	81.5%; Score 53; DB 7; Length 766;
Best Local Similarity	100.0%; Pred. No. 9.2;
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2 ALWWSPNG 9
Db	
	213 ALWWSPNG 220
RESULT 40	
ID	ADN39604
ID	ADN39604 standard; protein; 766 AA.
XX	AC ADN39604;
XX	17-JUN-2004 (first entry)
DE	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A204.
XX	Human; differential expression; cancer; angiogenic disorder;
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW	


```

AC ADO19398;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human PRO polypeptide #164.
XX
XX Human; PRO; immune related disorder; systemic lupus erythematosus;
XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
XX systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
XX autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
XX diabetes mellitus; renal disease; demyelinating disease;
XX central nervous system; peripheral nervous system;
XX demyelinating polyneuropathy; Guillain-Barre syndrome;
XX chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
XX WO2004043361-A2.
XX
XX 27-MAY-2004.
XX
XX 06-NOV-2003; 2003WO-US035268.
XX
XX 08-NOV-2002; 2002US-0425235P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
XX Wood WI, Wu TD;
XX
XX WPI; 2004-420067/39.
XX
XX N-PSDB; ADO19397.
XX
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
XX treating an immune related disorder such as systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX spondyloarthritis.
XX
XX Claim 7; SEQ ID NO 328; 1731pp; English.
XX
XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.
XX
XX Sequence 766 AA;
XX
XX Query Match 81.5%; Score 53; DB 8; Length 766;
XX Best Local Similarity 100.0%; Pred. No. 9.2;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 ALWWSPPNG 9
XX |||||
XX Db 213 ALWWSPPNG 220
XX
XX RESULT 44
XX ADO19400
XX ID ADO19400 standard; protein; 766 AA.
XX
XX AC ADO19400;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human PRO polypeptide #165.
XX
XX Human; PRO; immune related disorder; systemic lupus erythematosus;
XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
XX systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
XX autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
XX diabetes mellitus; renal disease; demyelinating disease;
XX central nervous system; peripheral nervous system;
XX demyelinating polyneuropathy; Guillain-Barre syndrome;
XX chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
XX WO2004043361-A2.
XX
XX 27-MAY-2004.
XX
XX 06-NOV-2003; 2003WO-US035268.
XX
XX 08-NOV-2002; 2002US-0425235P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
XX Wood WI, Wu TD;
XX
XX WPI; 2004-420067/39.
XX
XX N-PSDB; ADO19397.
XX
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
XX treating an immune related disorder such as systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX spondyloarthritis.
XX
XX Claim 7; SEQ ID NO 328; 1731pp; English.
XX
XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.
XX
XX Sequence 766 AA;
XX
XX Query Match 81.5%; Score 53; DB 8; Length 766;
XX Best Local Similarity 100.0%; Pred. No. 9.2;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 ALWWSPPNG 9
XX |||||
XX Db 213 ALWWSPPNG 220
XX
XX RESULT 45
XX ADO19806
XX ID ADO19806 standard; protein; 766 AA.
XX
XX AC ADO19806;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human PRO polypeptide #365.
XX
XX Human; PRO; immune related disorder; systemic lupus erythematosus;
XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
XX systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
XX autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
XX

```

KW diabetes mellitus; renal disease; demyelinating disease;
 KW central nervous system; peripheral nervous system;
 KW demyelinating polyneuropathy; Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polyneuropathy.
 XX Homo sapiens.
 XX WO2004043361-A2.
 XX 27-MAY-2004.
 XX 06-NOV-2003; 2003WO-US035268.
 XX 08-NOV-2002; 2002US-0425235P.
 XX (GETH) GENENTECH INC.
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX WPI; 2004-420067/39.
 XX N-PSDB; ADO19805.
 XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 XX Claim 7; SEQ ID NO 730; 1731pp; English.
 XX The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polypeptide of the
 CC invention.
 XX SQ Sequence 766 AA;
 Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWWSPNG 9
 Db 213 ALWWSPNG 220
 RESULT 46
 ADO71612
 ID ADO71612 standard; protein; 766 AA.
 AC ADO71612;
 XX 26-AUG-2004 (first entry)
 DT Amino acid sequence of a human CD26 protein.
 DE CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
 XX dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
 KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
 XX Homo sapiens.
 OS WO2004045497-A2.
 PN 03-JUN-2004.
 XX 15-MAY-2003; 2003WO-US015499.
 XX 17-MAY-2002; 2002US-0381606P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX 15-MAY-2003; 2003WO-US015499.
 XX 17-MAY-2002; 2002US-0381606P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Dang NH, Morimoto C;
 PI WPI; 2004-420511/39.
 XX N-PSDB; ADO71611, ADO71613.
 XX Use of a CD26 composition, and a chemotherapeutic and/or a
 PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
 PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
 PT regression or tumor necrosis.
 XX Claim 23; Page 151-153; 182pp; English.
 XX The specification describes a CD26 composition which, in conjunction with
 CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
 CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
 CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
 CC is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
 CC topoisomerase II inhibitor. The CD26 composition of the invention is
 CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
 CC a cell, killing a cancer cell, potentiating the effect of a
 CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
 CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
 CC inducing tumour regression or tumour necrosis. The CD26 composition is
 CC further useful for increasing topoisomerase II expression in a cell, for
 CC activating an antigen-presenting cell, or for potentiating immune
 CC responses of an animal. The present sequence represents a CD26 protein,
 CC and is encoded by vectors which are used to produce compositions of the
 CC invention.
 XX SQ Sequence 766 AA;
 Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWWSPNG 9
 Db 213 ALWWSPNG 220
 RESULT 47
 ADO71644
 ID ADO71644 standard; protein; 766 AA.
 XX ADO71644;
 AC 26-AUG-2004 (first entry)
 DT Amino acid sequence of a human CD26 protein.
 DE CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
 XX dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
 KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
 XX Homo sapiens.
 OS WO2004045497-A2.
 PN 03-JUN-2004.
 XX 15-MAY-2003; 2003WO-US015499.
 XX 17-MAY-2002; 2002US-0381606P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX

PI Dang NH, Morimoto C;
 XX WPI; 2004-420511/39.
 DR N-PSDB; ADO71643.
 XX
 PT Use of a CD26 composition, and a chemotherapeutic and/or a
 PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
 PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
 PT regression or tumor necrosis.
 XX
 PS Claim 23; Page 175-176; 182pp; English.
 XX
 CC The specification describes a CD26 composition which, in conjunction with
 CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
 CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
 CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
 CC is a dipeptidyl peptidase IV (DPP-IV). The chemotherapeutic agent is a
 CC topoisomerase II inhibitor. The CD26 composition of the invention is
 CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
 CC a cell, killing a cancer cell, potentiating the effect of a
 CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumor cell,
 CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
 CC inducing tumor regression or tumor necrosis. The CD26 composition is
 CC further useful for increasing topoisomerase II expression in a cell, for
 CC activating an antigen-presenting cell, or for potentiating immune
 CC responses of an animal. The present sequence represents a CD26 protein,
 CC and is encoded by vectors which are used to produce compositions of the
 CC invention.
 XX
 SQ Sequence 766 AA;
 Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 ALWMSPPNG 9
 |||||
 DB 213 ALWMSPPNG 220
 RESULT 48
 ADQ80365
 ID ADQ80365 standard; protein; 766 AA.
 XX
 AC ADQ80365;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Dipeptidylpeptidase IV protein.
 XX
 KW cystostatic; epidermal growth factor receptor modulator; identification;
 KW therapeutic response; cancer; EGFR; biomarker.
 XX
 OS Homo sapiens.
 XX
 PN WO2004063709-A2.
 XX
 PD 29-JUL-2004.
 XX
 PF 08-JAN-2004; 2004WO-US000368.
 XX
 PR 08-JAN-2003; 2003US-0438735P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Amler LC, Januario T;
 XX
 DR WPI; 2004-544114/52.
 DR N-PSDB; ADQ80241.
 XX
 PT Identifying a mammal that will respond therapeutically to a method of
 PT treating cancer comprises comparing the level of a biomarker in a mammal
 PT before and after exposure to an epidermal growth factor receptor (EGFR)

PT modulator.
 XX Disclosure; SEQ ID NO 137; 520pp; English.
 XX
 CC The invention relates to a method of identifying a mammal that will
 CC respond therapeutically to a method of treating cancer by administering
 CC an epidermal growth factor receptor (EGFR) modulator by comparing the
 CC level of a biomarker in a mammal before and after exposure to an EGFR
 CC modulator. The method comprises: (a) measuring, in the mammal, the level
 CC of at least one biomarker identified in the specification; (b) exposing
 CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
 CC level of the biomarker, where a difference in the level in step (c)
 CC compared to step (a) indicates that the mammal will respond
 CC therapeutically to the method of treating cancer. The method and
 CC biomarkers are useful for identifying a mammal that will respond
 CC therapeutically to a method of treating cancer by administering an
 CC epidermal growth factor receptor (EGFR) modulator. This sequence
 CC corresponds to one of the biomarkers whose levels of expression is
 CC measured in the method of the invention.
 XX
 SQ Sequence 766 AA;
 Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 ALWMSPPNG 9
 |||||
 DB 213 ALWMSPPNG 220
 RESULT 49
 ABM80355
 ID ABM80355 standard; protein; 766 AA.
 XX
 AC ABM80355;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Tumour-associated antigenic target (TAT) polypeptide PRO80881, SEQ:895.
 XX
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2004030615-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 29-SEP-2003; 2003WO-US028547.
 XX
 PR 02-OCT-2002; 2002US-0414971P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Wu TD, Zhang Z, Zhou Y;
 XX
 DR WPI; 2004-347921/32.
 DR N-PSDB; ACN37783.
 XX
 PT New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX
 PS Claim 12; SEQ ID NO 895; 7273pp; English.
 XX

CC The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention
 XX
 SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWSPNG 9
 Db 213 ALWSPNG 220
 |||||

RESULT 50
 ADP54458
 ID ADP54458 standard; protein; 766 AA.
 XX
 AC ADP54458;
 XX
 DT 18-NOV-2004 (first entry)
 DE Human PRO protein sequence SEQ ID NO:434.
 KW human; PRO; immune related disease; inflammatory immune response;
 KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
 KW antiaesthetic; antidiabetic; antiinflammatory; antipsoriatic;
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 KW vitucide; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2004039956-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 28-OCT-2003; 2003WO-US034381.
 XX
 PR 29-OCT-2002; 2002US-0422472P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX
 DR WPI; 2004-376182/35.
 DR N-PSDB; ADP54457.
 XX
 XX New PRO polynucleotides and polypeptides, useful in diagnosing
 PT and treating an immune related disease, e.g. systemic lupus
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 PT stimulating an immune response.
 XX

PS Claim 1; SEQ ID NO 434; 3009pp; English.
 XX
 CC The present invention describes an isolated PRO nucleic acid (I). Also
 CC described: (1) a vector comprising (I); (2) a host cell comprising the
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 CC antibody which specifically binds to a polypeptide of (4); (7) a
 CC composition of matter comprising a polypeptide of (4), an agonist or
 CC antagonist of the polypeptide or an antibody that binds to the
 CC polypeptide in combination with a carrier; (8) an article of manufacture
 CC comprising a container, a label on the container and a composition of
 CC matter of (7); (9) a method of treating an immune related disease in a
 CC mammal; (10) a method for determining the presence of a PRO polypeptide
 CC in a sample suspected of having the polypeptide; (11) a method of
 CC diagnosing an immune related disease or an inflammatory immune response
 CC in mammal; (12) a method of identifying a compound that inhibits or
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide
 CC ; and (13) a method of stimulating the immune response in a mammal. The
 CC PRO sequences have antiallergic, antianaemic, antiarthritic,
 CC antiaesthetic, antidiabetic, antiinflammatory, antipsoriatic,
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
 CC vitucide activities, and can be used in gene therapy. The nucleic acid
 CC (I) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO protein from the present invention.
 XX
 SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWSPNG 9
 Db 213 ALWSPNG 220
 |||||

RESULT 51
 ADU06688
 ID ADU06688 standard; protein; 766 AA.
 XX
 AC ADU06688;
 XX
 DT 27-JAN-2005 (first entry)
 DE Novel bronchial cancer-associated human protein SeqID914.
 KW bronchial cancer; cytostatic; tumour-associated protein;
 KW cancer detection; metastasis; tumour; human.
 XX
 OS Homo sapiens.
 XX
 PN DE10316701-A1.
 XX
 PD 04-NOV-2004.
 XX
 PF 09-APR-2003; 2003DE-01016701.
 XX
 PR 09-APR-2003; 2003DE-01016701.
 XX
 XX (HINZ/) HINZMANN B.
 PA (HERM/) HERMANN K.
 PA (CAST/) HEIDEN CASTANOS-VELEZ E.
 XX
 XX Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
 PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pillarsky C;
 XX WPI; 2004-786403/78.
 DR N-PSDB; ADU06201.

XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
 PT cancer and in screening for therapeutic and diagnostic agents.
 XX
 PS Claim 2; SEQ ID NO 914; 1381pp; German.
 XX
 CC This invention relates to a novel isolated nucleic acid associated with
 CC bronchial cancer comprising 489 defined sequences given in the
 CC specification. The invention may be useful for the production of
 CC compounds with a cytostatic activity through the inhibition of expression
 CC or activity of tumour-associated proteins. The novel DNA sequences and
 CC the proteins/peptides encoded by them are used for detecting bronchial
 CC cancer or determining the risk of developing it and to screen for
 CC specific binding partners of the DNA or protein sequences, where the
 CC binding partners are potentially useful as agents for treating or
 CC diagnosing bronchial cancer. The DNA or protein sequences can also be
 CC used for prognosis, detection of metastases and for secondary treatment
 CC (of tumours that have been stabilised or are no longer detectable).
 CC Detecting abnormal expression of the DNA sequences provides early
 CC diagnosis of bronchial cancers. The present sequence is that of a protein
 CC encoded by a novel bronchial cancer-associated human gene sequence of the
 CC invention.
 XX
 SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALWMSPNG 9
 |||||
 Db 213 ALWMSPNG 220

RESULT 52
 ADV25525
 ID ADV25525 standard; protein; 766 AA.
 XX
 AC ADV25525;
 XX
 DT 24-FEB-2005 (first entry)
 XX
 DE Human dipeptidyl-peptidase IV.
 XX
 KW Dipeptidyl-peptidase IV; DPP4; cardiovascular disease;
 KW dermatological disease; cancer; neoplasm; hematological disease;
 KW respiratory disease; gastrointestinal disease; liver disease;
 KW metabolic disorder; Cardiovascular-Gen.; Endocrine-Gen.;
 KW Antiinflammatory; Gastrointestinal-Gen.; Gynecological; Hepatotrophic;
 KW Neuroprotective; Cytostatic; Antiparkinsonian; Nootropic; Cardiant;
 KW Antiarrhythmic; Antiarteriosclerotic; Antianemic; Antidiabetic;
 KW Dermatological; Immunosuppressive; Muscular-Gen.; Antirheumatic;
 KW Antiarthritic; Antipsoriatic; Antiinfertility; Gene Therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2004104216-A2.
 XX
 PD 02-DEC-2004.
 XX
 PF 12-MAY-2004; 2004WO-EP005071.
 XX
 PR 21-MAY-2003; 2003EP-00011481.
 XX
 PA (FARB) BAYER HEALTHCARE AG.
 XX
 PI Golz S, Brueggemeier U, Summer H;
 XX
 DR WPI; 2004-834301/82.
 DR N-PSDB; ADV25524.
 XX
 PT Use of dipeptidylpeptidase IV (DPP4) polypeptides or polynucleotides for
 screening therapeutic agents or for diagnosing or treating diseases

PT associated with DPP4, e.g. cardiovascular, metabolic, inflammatory, or
 PT neurological disorders.
 XX
 PS Disclosure; SEQ ID NO 2; 128pp; English.

XX
 CC The present sequence is the protein sequence of human dipeptidyl-
 CC peptidase IV (DPP4). The invention relates to novel disease associations
 CC of DPP4 polypeptides and polynucleotides and to novel methods of
 CC screening for therapeutic agents for the treatment of cardiovascular
 CC disorders, dermatological disorders, cancer, hematological disorders,
 CC respiratory diseases, gastrointestinal and liver diseases, urological
 CC disorders and metabolic diseases. Pharmaceutical compositions are
 CC provided for treatment of these diseases and disorders and comprise a
 CC DPP4 polypeptide, a DPP4 polynucleotide, or regulators of DPP4 or
 CC modulators of DPP4 activity. The therapeutic agent is preferably a small
 CC molecule, an RNA molecule, an antisense oligonucleotide, a polypeptide,
 CC an antibody or a ribozyme. The invention also provides methods of
 CC diagnosing diseases and disorders associated with DPP4 by measuring the
 CC amount of a DPP4 polynucleotide in a sample and comparing it with the
 CC amount in a sample from a healthy and/or diseased mammal. The diseases
 CC and disorders include Parkinson's disease, dementia, Alzheimer's disease,
 CC myocardial infarction, arrhythmias, atherosclerosis, anemia, eosinophilic
 CC disorders, leukemia, pancreatitis, Crohn's disease, inflammatory bowel
 CC disease, diabetes, Cushing's syndrome, systemic lupus erythematosus,
 CC myasthenia gravis, rheumatoid arthritis, psoriasis, scleroderma, or
 CC infertility.

SQ Sequence 766 AA;

Query Match 81.5%; Score 53; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALWMSPNG 9
 |||||
 Db 213 ALWMSPNG 220

RESULT 53
 ADV15161
 ID ADV15161 standard; protein; 766 AA.
 XX
 AC ADV15161;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE PRO polypeptide SEQ ID NO 967.
 XX
 KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
 KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
 KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;
 KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
 KW Antiallergic; ds; gene; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN WO2005016962-A2.
 XX
 PD 24-FEB-2005.
 XX
 PF 11-AUG-2004; 2004WO-US026249.
 XX
 PR 11-AUG-2003; 2003US-0493546P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
 XX
 DR WPI; 2005-182330/19.
 XX
 PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
 PT treating an immune related disorder, e.g. systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

KW autism; nootropic; asperger syndrome; enzyme; dipeptidyl peptidase IV.
 XX Homo sapiens.
 OS
 XX US2005170333-A1.
 PN
 XX 04-AUG-2005.
 PD
 XX
 XX 03-FEB-2004; 2004US-00770712.
 PF
 XX
 XX 03-FEB-2004; 2004US-00770712.
 PR
 XX (VOJD/) VOJDANI A.
 PA
 XX Vojdani A;
 PI
 XX WPI; 2005-562713/57.
 DR
 XX
 XX Determining etiology of autistic spectrum disorder in patient, by
 PT determining level of infectious agent/toxic chemical/dietary protein
 PT derived antigen in samples of patient, comparing it with normal level of
 PT antigens of control subjects.
 XX
 XX Disclosure; SEQ ID NO 1; 89pp; English.
 PS
 XX
 XX The invention comprises a method of determining etiology of an autistic
 CC spectrum disorder in a patient. The method involves determining the level
 CC of an infectious agent, toxic chemical, or dietary protein derived
 CC antigen, or their antibodies in samples of patient, and comparing
 CC antigens/antibodies levels with normal levels of antigens/antibodies from
 CC control subjects. The method of the invention is useful for determining
 CC the etiology of an autistic spectrum disorder, such as autism, pervasive
 CC development disorder and Asperger's syndrome. The present amino acid
 CC sequence represents a human dipeptidyl peptidase IV enzyme that was used
 CC in the exemplification of the invention.
 XX
 XX Sequence 766 AA;
 SQ
 Query Match 81.5%; Score 53; DB 9; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWWSFNG 9
 |||||
 Db 213 ALWWSFNG 220
 RESULT 57
 AEB94223
 ID AEB94223 standard; protein; 766 AA.
 XX
 AC AEB94223;
 XX
 XX 06-OCT-2005 (first entry)
 DT
 XX
 XX CD26/dipeptidyl peptidase IV (DPPIV) SEQ ID NO:66.
 DE
 XX
 KW immune inhibition; fibroblast activation protein alpha dimer;
 KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
 KW immune disorder; neurological disease; autoimmune disease;
 KW immunosuppressive; graft versus host disease; transplant rejection;
 KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
 KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
 KW inflammation; respiratory disease; atherosclerosis; arteriosclerotic;
 KW cardiovascular disease; metabolic disorder; hashimoto's disease;
 KW antithyroid; endocrine disease; inflammatory bowel disease;
 KW antinflammatory; gastrointestinal-gen.; gastrointestinal disease;
 KW rheumatoid arthritis; autoimmune; multiple sclerosis; neuroprotective;
 KW autoimmune hepatitis; antiinflammatory; hepatotropic;
 KW systemic lupus erythematosus; dermatological; dermatological disease;
 KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
 KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
 KW psoriasis; antipsoriatic; graves disease; antithyroid;

KW respiratory syncytial virus infection; respiratory-gen.; virucide;
 KW CD26 dipeptidyl peptidase IV; DPPIV.
 XX Homo sapiens.
 OS
 XX WO2005071073-A1.
 PN
 XX 04-AUG-2005.
 PD
 XX
 XX 10-JAN-2005; 2005WO-US000709.
 PF
 XX
 XX 09-JAN-2004; 2004US-0535577P.
 PR
 XX (POIN-) POINT THERAPEUTICS INC.
 PA
 XX Mclean PA, Jones B, Miller GT, Jesson MI;
 PI
 XX WPI; 2005-564220/57.
 DR
 XX
 XX Down-regulating an immune response comprises administering to a subject
 PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
 PT amount effective to down-regulate an immune response.
 XX
 XX Disclosure; SEQ ID NO 66; 177pp; English.
 PS
 XX
 XX The invention relates to a method of down-regulating an immune response,
 CC which comprises administering to a subject a fibroblast activation
 CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
 CC an immune response. Also included are the following: a composition
 CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
 CC carrier, where the composition is sterile and lacks an adjuvant; a
 CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
 CC acceptable carrier, and a non-adjuvant second agent; a composition
 CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
 CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
 CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
 CC The method further comprises administering to the subject a second agent.
 CC The second agent is an anti-inflammatory agent, immunosuppressant, or
 CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
 CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
 CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
 CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
 CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
 CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
 CC substitution relative to wild type FAP alpha dimer. The amino acid
 CC substitution is present in the beta-propeller domain, the catalytic
 CC domain, or an N-linked glycosylation site and alters disulfide bond
 CC formation. The immune response is an especially an IL-1 mediated
 CC condition, abnormal immune response selected from inflammation,
 CC autoimmune disease, sepsis, graft versus host disease, transplant
 CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
 CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
 CC response is subsequent to an infection, such as an RSV infection. The
 CC autoimmune disease is selected from C, autoimmune thyroiditis, systemic
 CC lupus erythematosus (SLE), uveitis, hemolytic anemias, Graves' disease,
 CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
 CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
 CC sclerosis. The subject does not have cancer or a predisposition to
 CC cancer. The present sequence represents the amino acid sequence of human
 CC CD26/dipeptidyl peptidase IV (DPPIV).
 XX
 XX Sequence 766 AA;
 SQ
 Query Match 81.5%; Score 53; DB 9; Length 766;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWWSFNG 9
 |||||
 Db 213 ALWWSFNG 220
 RESULT 58

AAB11748
ID AAB11748 standard; protein; 767 AA.
AC AAB11748;
XX
XX 28-OCT-2000 (first entry)
DE Rat dipeptidyl peptidase IV (DPPIV).
XX
XX Dipeptidyl peptidase IV; DPPIV; rat; glomerulus;
KW F-16 monoclonal antibody; complement-mediated cytotoxicity; cell damage;
KW myocardial infarction; reperfusion injury; multiple sclerosis; nephritis;
KW neuropathy; arthritis; cardiant; neuroprotective; vasotropic;
KW antiarthritic.
XX
OS Rattus sp.
XX
XX
XX Key Location/Qualifiers
FH Region 35..767
FT /note= "This region is specifically claimed"
FT
XX JP2000143699-A.
PN 26-MAY-2000.
XX
XX 11-NOV-1998; 98JP-00321079.
XX
XX 11-NOV-1998; 98JP-00321079.
XX
XX (SHIO) SHIONOGI & CO LTD.
XX
XX WPI; 2000-445041/39.
XX
XX Monoclonal antibody specific for dipeptidyl peptidase, useful for
PT prevention and treatment of nephritis, arthritis, neuropathy and
PT myocardial infarction resulting from complement-dependent cell damage.
XX
XX Claim 2; Page 7-9; 11pp; Japanese.
XX
XX This sequence represents rat dipeptidyl peptidase IV (DPPIV). The
CC invention relates to the murine monoclonal antibody F-16 (FERMP-17016),
CC which suppresses complement-mediated cell damage by specifically binding
CC to DPPIV, thereby inhibiting complement activation. The invention also
CC relates to the F-16 hybridoma which secretes the monoclonal antibody and
CC a method of screening complement suppressive agents using the antibody
CC and DPPIV. The F-16 hybridoma was produced by fusion of a mouse myeloma
CC cell and a splenocyte obtained from a mouse inoculated with rat
CC glomerulus protein extract. The F-16 antibody was subsequently found to
CC bind to DPPIV. The F-16 antibody is useful for treating diseases caused
CC by complement-mediated cytotoxicity, such as myocardial infarction,
CC reperfusion injury, multiple sclerosis, nephritis, neuropathy and
CC arthritis. The antibody is also useful for screening agents with
CC potential complement suppression activity
XX
SQ Sequence 767 AA;
Query Match 81.5%; Score 53; DB 3; Length 767;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 ALWWSPPNG 9
| | | | | | | |
Db 211 ALWWSPPNG 218
RESULT 59
ABP56699
ID ABP56699 standard; protein; 767 AA.
XX
XX ABP56699;
XX
XX 27-MAR-2003 (first entry)
XX
DE Rat liver dipeptidyl peptidase IV (DPP4) protein SEQ ID NO:4.
XX
XX Tissue-specific endothelial membrane protein; antibacterial; virucide;
KW fungicide; antiparasitic; anticonvulsant; neuroleptic; cytostatic;
KW antiparkinsonian; nootropic; neuroprotective; antiasthmatic; infection;
KW antidiabetic; hypotensive; nephrotropic; antiarthritic; antiinflammatory;
KW gene therapy; epilepsy; schizophrenia; cancer; Parkinson's disease;
KW Alzheimer's disease; asthma; diabetes; hypertension; arthritis;
KW polycystic kidney disease; tissue-specific luminally expressed protein;
KW inflammatory bowel disease; enzyme.
XX
OS Rattus norvegicus.
XX
XX
XX Key Location/Qualifiers
FH Region 470
FT Misc-difference /note= "encoded by CAG"
FT
XX WO2002100336-A2.
PN 19-DEC-2002.
XX
XX 07-JUN-2002; 2002WO-US018185.
XX
XX 08-JUN-2001; 2001US-0297021P.
PR 12-JUN-2001; 2001US-0305117P.
XX
XX (TARG-) TARGET PROTEIN TECHNOLOGIES INC.
XX
XX Roben PW, Stevens AC;
PI
XX WPI; 2003-167367/16.
DR N-PSDB; ABZ22556.
XX
XX Delivering a therapeutic agent (e.g. immunosuppressant) to a specific
PT tissue, comprises administering an amount of a therapeutic complex
PT comprising a therapeutic group, a ligand that binds to a tissue-specific
PT protein, and a linker.
XX
XX Claim 24; Page 99-100; 123pp; English.
XX
XX The present invention describes a method (M1) for delivering a
CC therapeutic agent to a specific tissue, comprising administering an
CC amount of a therapeutic complex comprising a ligand which binds to a
CC tissue-specific luminally expressed protein, a therapeutic group, and a
CC linker which links the group to the ligand. Tissue-specific endothelial
CC membrane proteins from the present invention can have antibacterial,
CC virucide, fungicide, antiparasitic, anticonvulsant, neuroleptic,
CC cytotatic, antiparkinsonian, nootropic, neuroprotective, antiasthmatic,
CC antidiabetic, hypotensive, nephrotropic, antiarthritic and
CC antiinflammatory activities, and can be used in gene therapy. (M1) is
CC useful in targeting pharmaceuticals or other therapeutics to specific
CC tissues using tissue-specific endothelial membrane proteins. A
CC therapeutic complex may be used to treat or diagnose any disease for
CC which a tissue- or organ-specific treatment would be efficacious, such as
CC in cases of infections (e.g. bacterial, viral, fungal and parasitic),
CC epilepsy, schizophrenia, cancer, Parkinson's disease, Alzheimer's
CC disease, asthma, diabetes, hypertension, polycystic kidney disease,
CC arthritis, and inflammatory bowel disease. The present sequence
CC represents a rat liver dipeptidyl peptidase IV (DPP4), which is used in
CC an example from the present invention
XX
SQ Sequence 767 AA;
Query Match 81.5%; Score 53; DB 6; Length 767;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 ALWWSPPNG 9
| | | | | | | |
Db 211 ALWWSPPNG 218
RESULT 60
ABP56699
ID ABP56699 standard; protein; 767 AA.
XX
XX ABP56699;
XX
XX 27-MAR-2003 (first entry)
XX

ADD46932
 ID ADD46932 standard; protein; 767 AA.
 AC ADD46932;
 XX
 XX
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 XX
 XX
 DE Rat Protein AAA41096, SEQ ID NO 12618.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 OS Unidentified.
 XX
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; AAA41096.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Example 1; Page; 1017pp; English.
 XX

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (described in Table 3 of the specification) which is differentially expressed during pain.
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 767 AA;
 Query Match 81.5%; Score 53; DB 7; Length 767;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

This page blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:19:28 ; Search time 15.7895 Seconds
(without alignments)
60.937 Million cell updates/sec

Title: US-10-774-242A-4
Perfect score: 65
Sequence: 1 YALMWSNGK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : PIR 80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	759	138593	fibroblast activat
2	53	81.5	766	1 CDH926	dipeptidyl-peptida
3	53	81.5	792	1 A39914	dipeptidyl-peptida
4	47	72.3	760	1 S23752	dipeptidyl-peptida
5	47	72.3	829	2 T19514	hypothetical prote
6	46	70.8	509	2 D86146	F22L4.11 protein -
7	45	69.2	250	2 AC2988	3-demethylubiquino
8	45	69.2	250	2 E98295	dihydroxypolypreny
9	45	69.2	320	2 T26259	hypothetical prote
10	45	69.2	743	2 T37700	probable dipeptidy
11	44	67.7	699	2 H64118	4-alpha-glucanotra
12	43	66.2	237	2 G75316	probable lipase -
13	43	66.2	582	2 I48673	matrix metalloprot
14	43	66.2	582	2 I38028	matrix metalloprot
15	43	66.2	582	2 I84471	matrix metalloprot
16	42	64.6	248	2 AG3275	hexaprenyldihydrox
17	42	64.6	455	2 A72050	Glpt/Pgpt/Uhpt fam
18	42	64.6	455	2 F86573	hexosophosphate tra
19	42	64.6	818	1 A30107	dipeptidyl amino
20	41	63.1	246	2 T01868	probable glucose-6
21	41	63.1	252	2 E87353	3-demethylubiquino
22	41	63.1	669	2 I38029	matrix metalloprot
23	41	63.1	785	2 T01541	hypothetical prote
24	41	63.1	803	2 I68600	dipeptidyl amino
25	41	63.1	803	2 A41793	dipeptidyl amino
26	41	63.1	855	2 I54331	dipeptidyl amino
27	40	61.5	351	2 JQ2166	spindle body prote
28	40	61.5	352	2 JC5185	enhancing factor p
29	40	61.5	421	2 T18825	hypothetical prote

30	40	61.5	466	2 T30040	hypothetical prote
31	40	61.5	539	2 AR0998	phosphoenolpyruvat
32	40	61.5	724	2 T20450	hypothetical prote
33	40	61.5	724	2 G70928	probable 4-ALPHA-G
34	40	61.5	2057	2 F90109	splicing factor Pr
35	40	61.5	2329	2 S44625	C50C3.6 protein -
36	40	61.5	2352	2 T06077	splicing factor PR
37	40	61.5	2359	2 B96832	hypothetical prote
38	40	61.5	2363	2 T38841	probable pre-mRNA
39	40	61.5	2413	2 S34670	splicing factor PR
40	39	60.0	168	2 T35629	hypothetical prote
41	39	60.0	185	2 G87558	conserved hypothet
42	39	60.0	223	2 T19589	hypothetical prote
43	39	60.0	301	2 B84282	8-oxoguanine DNA g
44	39	60.0	305	2 F46110	hypothetical prote
45	39	60.0	321	2 T08544	hypothetical prote
46	39	60.0	350	2 S22366	phenylalanine-tRNA
47	39	60.0	376	2 T40673	homoserine dehydro
48	39	60.0	409	1 SVZJAL	5-aminolevulinate
49	39	60.0	465	2 A32569	U4/U6 snRNP 52K pr
50	39	60.0	666	2 T40172	Glucose inhibited
51	39	60.0	766	1 S61694	flocculation suppr
52	39	60.0	802	2 C86427	hypothetical prote
53	39	60.0	879	2 T49796	probable sepB prot
54	39	60.0	1002	2 T43236	carboxypeptidase C
55	39	60.0	1072	2 AD1280	SNF2-type helicase
56	39	60.0	1072	2 AD1643	SNF2-type helicase
57	39	60.0	1223	2 E88451	PRP8 protein homol
58	39	60.0	2403	2 T30875	variant-specific s
59	39	60.0	2924	2 T18378	genome polyprotein
60	39	60.0	3473	1 A46112	polyprotein - rice
61	39	60.0	3473	2 S27927	superoxide dismuta
62	38.5	59.2	137	2 S52366	superoxide dismuta
63	38.5	59.2	137	2 S52366	superoxide dismuta
64	38.5	59.2	137	2 S52366	superoxide dismuta
65	38.5	59.2	207	2 JC4351	superoxide dismuta
66	38.5	59.2	207	2 S60669	superoxide dismuta
67	38.5	59.2	207	2 S05599	conserved hypothet
68	38	58.5	204	2 H82283	hypothetical prote
69	38	58.5	204	2 G64006	leucyl/phenylalan
70	38	58.5	226	2 G83318	probable hydrolase
71	38	58.5	232	2 B82987	leucyl, phenylalan
72	38	58.5	234	2 H85613	leucyl, phenylalan
73	38	58.5	234	2 AI0610	leucyl, phenylalan
74	38	58.5	234	2 A36888	leucyl, phenylalan
75	38	58.5	234	2 B90750	probable leucyl/ph
76	38	58.5	236	2 AE0167	protein transferas
77	38	58.5	243	2 D82681	leucyl/phenylalan
78	38	58.5	243	2 G82163	probable exported
79	38	58.5	255	2 AC0384	secretion protein
80	38	58.5	279	1 S32869	protein F12A21.25
81	38	58.5	296	2 F96699	hypothetical prote
82	38	58.5	373	2 T34699	hypothetical prote
83	38	58.5	403	2 E83366	virulence-mediati
84	38	58.5	444	2 JC4348	hypothetical prote
85	38	58.5	485	2 T25684	erythropoietin rec
86	38	58.5	507	1 A46713	hypothetical prote
87	38	58.5	532	2 T02539	4-alpha-glucanotra
88	38	58.5	726	2 B82511	hypothetical prote
89	38	58.5	766	2 AF2272	DNA uptake protein
90	38	58.5	836	2 C82726	conserved hypothet
91	38	58.5	906	2 G90281	hypothetical prote
92	38	58.5	951	2 T47324	superoxide dismuta
93	37.5	57.7	137	2 S52362	superoxide dismuta
94	37.5	57.7	137	2 S52366	superoxide dismuta
95	37.5	57.7	137	2 S52366	superoxide dismuta
96	37.5	57.7	137	2 S52376	superoxide dismuta
97	37.5	57.7	207	2 S15205	hypothetical prote
98	37	56.9	120	2 C82909	lysozyme (EC 3.2.1
99	37	56.9	129	1 LZDK3	lysozyme (EC 3.2.1
100	37	56.9	147	1 LZDK	conserved hypothet
101	37	56.9	154	2 G56031	hypothetical prote
102	37	56.9	209	2 C85295	hypothetical prote

103	37	56.9	210	2	T05792	hypothetical prote	176	36	55.4	858	1	IJRTNC	neural cell adhesi
104	37	56.9	214	2	D83171	conserved hypothet	177	36	55.4	888	1	GNLJHD	pol polyprotein -
105	37	56.9	257	2	T22764	hypothetical prote	178	36	55.4	921	2	E81076	TonB-dependent rec
106	37	56.9	264	2	AD0477	probable membrane	179	36	55.4	922	2	G81865	probable outer mem
107	37	56.9	302	2	B84979	sulfate adenylitr	180	36	55.4	1091	1	IUCHNL	neural cell adhesi
108	37	56.9	341	2	JCS855	polyketide synthas	181	36	55.4	1115	1	IJMSNL	neural cell adhesi
109	37	56.9	462	2	T50168	probable U4/U6 sma	182	36	55.4	1215	2	T11660	probable RNA-dir
110	37	56.9	507	1	A32385	erythropoietin rec	183	36	55.4	1290	2	A57190	ebnerin precursor
111	37	56.9	508	1	ZUHUR	erythropoietin rec	184	36	55.4	1390	2	T18883	hypothetical prote
112	37	56.9	509	2	A72314	hypothetical prote	185	36	55.4	1486	1	B40333	collagen alpha 1(I
113	37	56.9	519	2	S75570	apolipoprotein N-a	186	36	55.4	1492	2	A40333	collagen alpha 1(I
114	37	56.9	544	2	AF2248	hypothetical prote	187	36	55.4	1492	2	T42721	CRP-ductin-alpha p
115	37	56.9	548	2	T11732	ABC transporter At	188	35	55.4	2083	2	D96633	hypothetical prote
116	37	56.9	586	2	T45210	hypothetical prote	189	35	54.6	664	2	D96633	hypothetical prote
117	37	56.9	589	2	F97735	hypothetical prote	190	35	53.8	71	2	G84978	probable ribosomal
118	37	56.9	592	2	S56681	beta-fructofuranos	191	35	53.8	87	1	QQQYVW	conserved hypothet
119	37	56.9	594	2	E96667	unknown protein, 6	192	35	53.8	103	2	AF0856	hypothetical prote
120	37	56.9	597	2	G84825	probable CCH-type	193	35	53.8	103	2	B91079	hypothetical prote
121	37	56.9	600	2	T48433	hypothetical prote	194	35	53.8	103	2	H65055	hypothetical prote
122	37	56.9	611	2	D84423	probable WD-40-rep	195	35	53.8	103	2	C85924	hypothetical prote
123	37	56.9	662	2	D86320	hypothetical prote	196	35	53.8	151	2	S04123	hypothetical prote
124	37	56.9	680	2	T45736	hypothetical prote	197	35	53.8	151	2	B95345	hypothetical prote
125	37	56.9	779	2	T25173	hypothetical prote	198	35	53.8	151	2	F95339	hypothetical prote
126	37	56.9	787	2	T00798	hypothetical prote	199	35	53.8	169	2	G83736	hypothetical prote
127	37	56.9	799	2	T25174	hypothetical prote	200	35	53.8	176	2	A86815	transcription regu
128	37	56.9	815	2	T00538	probable serine pr	201	35	53.8	188	2	T37183	hypothetical prote
129	37	56.9	836	2	S54152	seps protein - Eme	202	35	53.8	190	2	S48101	xyloglucan endo-1,
130	37	56.9	931	2	A49737	dipeptidyl aminope	203	35	53.8	200	2	G90298	hypothetical prote
131	37	56.9	1142	2	E96519	probable reverse t	204	35	53.8	211	2	T49671	hypothetical prote
132	37	56.9	1163	2	E71481	probable awf/snf h	205	35	53.8	215	2	A95313	hypothetical prote
133	37	56.9	1166	2	G86596	SWI/SNF family hel	206	35	53.8	252	1	G71667	probable 3-dimethy
134	37	56.9	1166	2	T2027	swi/snf family hel	207	35	53.8	253	2	T36718	probable integral
135	37	56.9	1534	2	S59604	DNA (cytosine-5)-	208	35	53.8	259	2	C98356	conserved hypothet
136	37	56.9	2228	2	T14029	variant-specific s	209	35	53.8	259	2	S73463	hypothetical prote
137	36	55.4	129	2	AG2003	hypothetical prote	210	35	53.8	269	2	T36001	probable integral
138	36	55.4	136	2	F75309	hypothetical prote	211	35	53.8	273	2	G98342	hypothetical prote
139	36	55.4	181	2	S65230	hypothetical prote	212	35	53.8	273	2	A42939	conserved hypothet
140	36	55.4	185	2	T08543	hypothetical prote	213	35	53.8	284	2	G72662	hypothetical prote
141	36	55.4	221	2	G83964	hypothetical prote	214	35	53.8	293	2	G57750	RNA polymerase sig
142	36	55.4	246	2	H82553	3-demethylubiquino	215	35	53.8	293	2	B71686	RNA polymerase sig
143	36	55.4	293	2	T48975	xyloglucan endo-tr	216	35	53.8	293	2	E85363	hypothetical prote
144	36	55.4	323	2	A61396	collagen alpha 1(I	217	35	53.8	295	2	S48102	xyloglucan endo-1,
145	36	55.4	328	2	A48971	IS1380 polypeptide	218	35	53.8	296	2	JC7268	paired-type homeod
146	36	55.4	343	2	T22047	hypothetical prote	219	35	53.8	298	2	H87632	heat shock sigma f
147	36	55.4	365	2	A72574	probable acyl carr	220	35	53.8	300	2	S60164	transcription init
148	36	55.4	381	2	S68165	glycerate dehydrog	221	35	53.8	300	2	AC2877	RNA polymerase sig
149	36	55.4	382	1	DEKVG	glycerate dehydrog	222	35	53.8	300	2	E97653	transcription init
150	36	55.4	386	2	S68164	glycerate dehydrog	223	35	53.8	302	2	S60169	transcription init
151	36	55.4	405	2	B36340	alpha(1,3)-fucosyl	224	35	53.8	303	2	AD3299	RNA polymerase sig
152	36	55.4	414	2	AB0837	probable ferric en	225	35	53.8	305	2	F71904	probable outer mem
153	36	55.4	443	2	A71686	tolb protein precu	226	35	53.8	315	2	T16359	hypothetical prote
154	36	55.4	444	2	F97750	tolb protein precu	227	35	53.8	319	2	A70715	hypothetical prote
155	36	55.4	507	1	FWCNBA	beta-globulin A pr	228	35	53.8	320	2	A40660	outer membrane pro
156	36	55.4	533	2	S47582	high-affinity pota	229	35	53.8	325	2	T31929	hypothetical prote
157	36	55.4	546	2	AB0916	ubiquinone biosynt	230	35	53.8	339	2	S46409	hypothetical prote
158	36	55.4	546	2	C86071	Probable ubiquinol	231	35	53.8	349	2	F69205	3-hydroxy-3-methyl
159	36	55.4	546	2	D65188	Probable ubiquinol	232	35	53.8	354	2	JC4212	nitrilase (EC 3.5.
160	36	55.4	546	2	E91224	conserved hypothet	233	35	53.8	355	2	C90144	terminal oxidase,
161	36	55.4	593	2	H71283	conserved hypothet	234	35	53.8	363	2	T36408	probable esterase
162	36	55.4	602	2	G75278	N-acetylmuramoyl-L	235	35	53.8	402	2	D82980	hypothetical prote
163	36	55.4	605	2	E71253	alanine-tRNA ligas	236	35	53.8	404	2	T02282	hypothetical prote
164	36	55.4	633	2	T05005	hypothetical prote	237	35	53.8	432	2	E87508	outer membrane pro
165	36	55.4	649	2	E98215	hypothetical prote	238	35	53.8	467	2	AG0546	probable terminal
166	36	55.4	655	2	AD3071	hypothetical prote	239	35	53.8	477	2	AE0608	conserved hypothet
167	36	55.4	679	2	AC0333	probable membrane	240	35	53.8	482	2	AF1120	conserved hypothet
168	36	55.4	694	1	C65137	4-alpha-glucanotra	241	35	53.8	494	2	A11480	probable cytochrom
169	36	55.4	694	2	A86007	4-alpha-glucanotra	242	35	53.8	508	2	T16980	transcription init
170	36	55.4	694	2	B91161	4-alpha-glucanotra	243	35	53.8	532	2	S78176	transcription init
171	36	55.4	725	1	IJMSNG	neural cell adhesi	244	35	53.8	573	2	F81313	peptidase (M3 fami
172	36	55.4	737	2	T16737	hypothetical prote	245	35	53.8	578	1	F64578	oligoendopeptidase
173	36	55.4	761	1	IJHUNG	neural cell adhesi	246	35	53.8	578	2	B71934	oligoendopeptidase - H
174	36	55.4	793	2	T41703	dipeptidyl aminope	247	35	53.8	606	2	G72282	hypothetical prote
175	36	55.4	853	1	IJBONC	neural cell adhesi	248	35	53.8	609	2	JC6134	DNA-directed RNA p

249	35	53.8	631	2	T35234	probable secreted	322	34	52.3	379	2	T19588	hypothetical prote
250	35	53.8	643	2	H88712	protein C17H12.1 (323	34	52.3	386	2	C83436	probable MPS metab
251	35	53.8	698	2	AC0016	4-alpha-glucanotra	324	34	52.3	392	2	C82769	phage-related cont
252	35	53.8	708	1	M2XR4D	structural protein	325	34	52.3	399	2	S34681	hypothetical prote
253	35	53.8	708	1	M2XR1L	structural protein	326	34	52.3	409	2	B70846	probable icdl prot
254	35	53.8	708	1	M2XR2J	structural protein	327	34	52.3	412	2	S28423	isocitrate dehydro
255	35	53.8	708	1	M2XR2L	structural protein	328	34	52.3	415	2	S65065	isocitrate dehydro
256	35	53.8	709	1	M2XR3D	structural protein	329	34	52.3	416	2	S47013	isocitrate dehydro
257	35	53.8	709	1	M2XR3D	structural protein	330	34	52.3	416	2	C47017	probable transcrip
258	35	53.8	796	2	G64570	probable s/n-oxide r	331	34	52.3	416	2	AF2017	RNA polymerase sig
259	35	53.8	806	2	G95362	probable (imported	332	34	52.3	426	2	F81653	isocitrate dehydro
260	35	53.8	849	2	T43342	fzr related protai	333	34	52.3	428	1	DCBYIS	isocitrate dehydro
261	35	53.8	886	2	A87093	valyl-tRNA synthas	334	34	52.3	428	2	E95906	probable solute-bi
262	35	53.8	952	2	AG1209	B. subtilis YfhO p	335	34	52.3	430	2	F82625	TolB protein precu
263	35	53.8	1021	2	T10748	mannan endo-1,4-be	336	34	52.3	430	2	AG0209	probable sugar-bin
264	35	53.8	1217	2	S52714	sericin1B - silkwo	337	34	52.3	431	2	A72037	TolB protein, prob
265	35	53.8	1635	2	T14075	chitinase (EC 3.2.	338	34	52.3	431	2	D86588	macromolecule tran
266	35	53.8	1638	2	T25352	hypothetical prote	339	34	52.3	431	2	D71495	probable polysacch
267	34.5	53.1	142	2	H69870	hypothetical prote	340	34	52.3	433	2	T09619	hypothetical prote
268	34.5	53.1	193	2	G86901	hypothetical prote	341	34	52.3	435	2	T19733	probable carboxyl-
269	34.5	53.1	293	2	T50244	probable ammonium	342	34	52.3	436	2	B83005	probable isocitrat
270	34	52.3	512	2	S58987	H+-transporting tw	343	34	52.3	438	2	T07402	hypothetical prote
271	34	52.3	93	2	F84175	hypothetical prote	344	34	52.3	440	2	T51278	unknown protein li
272	34	52.3	100	2	F95369	hypothetical prote	345	34	52.3	443	2	H95542	isocitrate dehydro
273	34	52.3	104	2	D72542	hypothetical prote	346	34	52.3	451	2	S33612	isocitrate dehydro
274	34	52.3	107	2	JQ0143	hypothetical 12.4k	347	34	52.3	451	2	F87659	putrescine/ornithi
275	34	52.3	108	2	S75540	hypothetical prote	348	34	52.3	452	2	G95306	epoxide hydrolase
276	34	52.3	112	2	B30503	Ig gamma-2a chain	349	34	52.3	455	1	A26081	endo-xylanase homo
277	34	52.3	137	2	FQ0502	hypothetical prote	350	34	52.3	462	2	T17480	isocitrate dehydro
278	34	52.3	157	2	G82214	hypothetical prote	351	34	52.3	465	2	T48632	hypothetical prote
279	34	52.3	160	2	S60172	potassium channel	352	34	52.3	471	2	A86303	hypothetical prote
280	34	52.3	160	2	T16043	hypothetical prote	353	34	52.3	474	2	T27297	conserved hypotet
281	34	52.3	182	2	H95055	conserved hypotet	354	34	52.3	478	2	F82175	transcription fact
282	34	52.3	182	2	E97925	hypothetical prote	355	34	52.3	479	1	A31753	transcription fact
283	34	52.3	192	2	T27838	hypothetical prote	356	34	52.3	483	1	H90210	hypothetical prote
284	34	52.3	193	2	A36672	transcription init	357	34	52.3	485	1	S22543	transcription fact
285	34	52.3	197	2	E83754	hypothetical prote	358	34	52.3	485	1	S22543	transcription fact
286	34	52.3	220	2	A70526	hypothetical prote	359	34	52.3	489	2	H84106	teichuronopeptide
287	34	52.3	227	2	T22144	hypothetical prote	360	34	52.3	492	2	A70630	probable lpgm prot
288	34	52.3	231	2	AG2277	hypothetical prote	361	34	52.3	502	2	T41148	trp-asp repeat con
289	34	52.3	246	2	T10127	transcription init.	362	34	52.3	514	1	P1WL47	l1 protein - human
290	34	52.3	256	2	C97062	potassium channel	363	34	52.3	516	1	P1WL5	l1 protein - human
291	34	52.3	263	2	JH0635	type II site-speci	364	34	52.3	520	2	F69470	dipteric ABC tran
292	34	52.3	272	2	T01480	hypothetical prote	365	34	52.3	525	1	P1WLBS	l1 protein - human
293	34	52.3	281	2	A55849	ketopantoate reduc	366	34	52.3	528	2	D85048	hypothetical prote
294	34	52.3	282	2	B72859	deoxyhypusine synt	367	34	52.3	539	2	AG0017	phosphoenolpyruvat
295	34	52.3	283	2	AC2780	tolB protein (impo	368	34	52.3	543	2	T37570	WD repeat protein li
296	34	52.3	293	2	E97559	tolB protein precu	369	34	52.3	548	2	C86456	probable transpos
297	34	52.3	293	2	E83513	probable outer mem	370	34	52.3	550	2	B70583	probable integral
298	34	52.3	295	2	A40587	hypothetical prote	371	34	52.3	589	2	T34878	two-component sens
299	34	52.3	295	2	AC2451	transcription init	372	34	52.3	589	2	H69691	hypothetical prote
300	34	52.3	296	2	S30249	transcription init	373	34	52.3	604	2	T21731	hypothetical prote
301	34	52.3	303	2	AB0556	ApbA (imported) -	374	34	52.3	604	2	T02684	ABC transporter, p
302	34	52.3	305	2	S77033	hypothetical prote	375	34	52.3	611	2	T15410	hypothetical prote
303	34	52.3	309	2	AC1351	ferrochelatase hom	376	34	52.3	622	2	A64494	hypothetical prote
304	34	52.3	309	2	AF1721	ferrochelatase hom	377	34	52.3	686	2	T06700	hypothetical prote
305	34	52.3	310	2	C70426	RNA polymerase sig	378	34	52.3	689	2	T35882	hypothetical prote
306	34	52.3	318	2	H83298	conserved hypotet	379	34	52.3	692	2	H70362	glycogen phosphory
307	34	52.3	329	2	A64669	probable aldo-keto	380	34	52.3	693	2	D90441	hypothetical prote
308	34	52.3	332	2	B47017	probable transcrip	381	34	52.3	726	2	T34638	hypothetical prote
309	34	52.3	332	2	AD2541	transcription init	382	34	52.3	738	2	AB7516	dipteric peptida
310	34	52.3	333	2	A47536	gene WNT3 protein	383	34	52.3	753	2	AH0097	probable biotin su
311	34	52.3	338	2	T43440	hypothetical prote	384	34	52.3	757	2	T16149	hypothetical prote
312	34	52.3	350	2	AG3326	serine/threonine p	385	34	52.3	768	2	JC5564	cellulose oxidase
313	34	52.3	351	2	B34840	heterogeneous ribo	386	34	52.3	788	2	T30510	viral enhancing fa
314	34	52.3	352	2	A48828	wingless homolog X	387	34	52.3	798	2	AI2053	competence protein
315	34	52.3	352	2	A39532	Wnt-3A protein - m	388	34	52.3	864	2	JC4624	alpha-glucosidase
316	34	52.3	353	2	AH2024	hypothetical prote	389	34	52.3	889	2	S51465	hypothetical prote
317	34	52.3	355	2	A35503	Wnt-3 protein - mo	390	34	52.3	932	2	T40216	probable histone t
318	34	52.3	356	2	T10888	iron utilization p	391	34	52.3	949	2	E75352	glycine cleavage a
319	34	52.3	356	2	E64048	nucleotide-binding	392	34	52.3	952	2	T48510	probable histone t
320	34	52.3	365	2	A34840	heterogeneous ribo	393	34	52.3	995	2	T32466	MTB like protein -
321	34	52.3	378	2	H70548	hypothetical prote	394	34	52.3	999	2	T27628	hypothetical prote

395 34 52.3 1009 2 T16604 hypothetical prote
396 34 52.3 1032 2 D95177 SNF2 family protei
397 34 52.3 1032 2 P98043 SWI/SNF family ATP
398 34 52.3 1034 2 B86880 hypothetical prote
399 34 52.3 1062 2 D95840 hypothetical prote
400 34 52.3 1119 2 T15842 DNA repair protein
401 34 52.3 1169 1 S64859 mannosyltransferas
402 34 52.3 1213 2 E69255 hypothetical prote
403 34 52.3 1227 2 A86245 hypothetical prote
404 34 52.3 1272 2 C96637 hypothetical prote
405 34 52.3 1280 2 T29021 hypothetical prote
406 34 52.3 1389 2 H84914 probable WD-40 rep
407 34 52.3 1607 2 T02837 long chain fatty a
408 34 52.3 1817 2 T34249 hypothetical prote
409 34 52.3 2135 2 T14602 variant-specific s
410 34 52.3 2212 2 T28157 erythrocyte membra
411 34 52.3 2647 2 T28161 hypothetical prote
412 34 52.3 2706 2 T28155 variant-specific s
413 34 52.3 3005 2 T08841 polypeptide - dour
414 34 52.3 3006 2 T28625 variant-specific s
415 33.5 51.5 137 2 S52370 superoxide dismuta
416 33.5 51.5 137 2 S52372 superoxide dismuta
417 33.5 51.5 179 2 T20206 hypothetical prote
418 33.5 51.5 379 2 T35334 3-oxoacyl-acyl car
419 33.5 51.5 480 2 T23608 hypothetical prote
420 33.5 51.5 506 2 T40396 glucosyltransferas
421 33.5 51.5 536 2 G85488 probable transport
422 33.5 51.5 536 2 G90637 probable transport
423 33.5 51.5 555 2 E87576 choline dehydrogen
424 33.5 51.5 1047 2 AP1912 hypothetical prote
425 33.5 51.5 1083 2 H86921 probable arabinosy
426 33 50.8 61 1 WMXF12 P12 attachment pro
427 33 50.8 83 2 T17961 hypothetical prote
428 33 50.8 87 1 D75153 probable ribosomal
429 33 50.8 110 2 T30125 hypothetical prote
430 33 50.8 118 2 T05812 hypothetical prote
431 33 50.8 124 2 T45917 hypothetical prote
432 33 50.8 127 2 E87662 hypothetical prote
433 33 50.8 130 2 H96507 hypothetical prote
434 33 50.8 130 2 T47309 hypothetical prote
435 33 50.8 131 2 A84491 hypothetical prote
436 33 50.8 146 2 T47310 hypothetical prote
437 33 50.8 151 2 S37013 transposase (Clone
438 33 50.8 160 2 AF2171 hypothetical prote
439 33 50.8 175 2 D86471 hypothetical prote
440 33 50.8 176 2 T47361 hypothetical prote
441 33 50.8 181 2 T15378 hypothetical prote
442 33 50.8 182 2 C90429 hypothetical prote
443 33 50.8 185 1 WMEWSC P4 protein - south
444 33 50.8 190 2 E86214 protein F24B9.1 [i
445 33 50.8 196 2 T29016 hypothetical prote
446 33 50.8 199 2 A10736 hypothetical prote
447 33 50.8 205 2 T37103 hypothetical prote
448 33 50.8 209 2 S44297 hypothetical prote
449 33 50.8 209 2 T08451 hypothetical prote
450 33 50.8 210 2 T04233 pathogenesis-relat
451 33 50.8 215 2 A99734 probable tail asse
452 33 50.8 215 2 C85584 probable tail comp
453 33 50.8 228 2 S46965 microfilarial shea
454 33 50.8 237 2 G46964 microfilarial shea
455 33 50.8 245 2 G81818 probable periplasm
456 33 50.8 245 2 E81057 conserved hypothet
457 33 50.8 246 2 AC3292 hypothetical prote
458 33 50.8 249 2 H97251 hypothetical prote
459 33 50.8 254 2 H83353 probable permease
460 33 50.8 255 2 B87595 transcriptions regu
461 33 50.8 255 2 G75157 abc transporter, A
462 33 50.8 255 2 B71197 hypothetical prote
463 33 50.8 257 2 C70635 hypothetical prote
464 33 50.8 273 2 T42928 immediate-early pr
465 33 50.8 275 2 S77462 hypothetical prote
466 33 50.8 276 2 B86370 32.3K hypothetical
467 33 50.8 280 2 T22048 hypothetical prote

468 33 50.8 281 2 E87703 conserved hypothet
469 33 50.8 283 2 T11063 NADH2 dehydrogenas
470 33 50.8 283 2 A12345 formamido-pyrimidin
471 33 50.8 286 2 E69516 hypothetical prote
472 33 50.8 290 2 T16212 hypothetical prote
473 33 50.8 299 2 F84785 probable xyloglucan
474 33 50.8 310 2 T43158 probable GTP-bindin
475 33 50.8 314 2 T43299 probable GTP-bindin
476 33 50.8 318 2 T20063 hypothetical prote
477 33 50.8 319 2 E87515 hypothetical prote
478 33 50.8 328 2 A91012 hypothetical prote
479 33 50.8 328 2 C85856 hypothetical prote
480 33 50.8 328 2 D64986 yeast protein - Esc
481 33 50.8 328 2 B86280 NADH2 dehydrogenas
482 33 50.8 332 2 S63660 conserved hypothet
483 33 50.8 332 2 F83600 malic acid transpo
484 33 50.8 342 2 B64395 probable PPE prote
485 33 50.8 346 2 H70874 hypothetical prote
486 33 50.8 347 2 B70710 hypothetical prote
487 33 50.8 347 2 H64371 malic acid transpo
488 33 50.8 349 2 A44507 licheninase (EC 3.
489 33 50.8 352 2 T36719 probable integral.
490 33 50.8 355 2 H87413 hypothetical prote
491 33 50.8 358 2 S45911 hypothetical prote
492 33 50.8 363 2 H98101 conserved hypothet
493 33 50.8 363 2 G95237 conserved hypothet
494 33 50.8 363 2 A95273 protein F12M16.15
495 33 50.8 365 2 T20958 hypothetical prote
496 33 50.8 369 2 S75038 hypothetical prote
497 33 50.8 376 2 S04497 surface antigen PA
498 33 50.8 396 2 AD0354 dihydropteridine r
499 33 50.8 400 2 JC4591 alpha-1,3 fucosylt
500 33 50.8 401 1 T21288 phosphoprotein pho

ALIGNMENTS

RESULT 1

I38593
I38593 fibroblast activation protein-alpha - human
N:Alternate names: FAP-alpha
C:Species: Homo sapiens (man)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C:Accession: I38593
R:Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncasi, M.P.; Healey, J.; O.
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A:Title: Molecular cloning of fibroblast activation protein alpha, a member of the serin
A:Reference number: I38593; MUID:94261645; PMID:7911242
A:Accession: I38593
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-759 <RES>
A:Cross-references: UNIPARC:UPI0000175B7D; EMBL:U09278; NID:g507749; PID:g507750
C:Genetics:
A:Gene: GDB:FAP
A:Cross-references: GDB:374184; OMIM:600403
A:Map position: 2q23-2q23
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: transmembrane protein

Query Match 100.0%; Score 65; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YALWWSNPK 10
Db 210 YALWWSNPK 219
RESULT 2
CDHU26 dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
 C;Species: Homo sapiens (man)
 C;Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004
 C;Accession: S24313; B42408; A42408; B61136; S59510; I56154; S59857; S15520
 R;Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.
 Biochim. Biophys. Acta 1131, 333-336, 1992
 A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease
 A;Reference number: S24313; MUID:92329551; PMID:1352704
 A;Accession: S24313
 A;Molecule type: mRNA
 A;Residues: 1-61, 1-8-766 <MS>
 A;Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AE80; EMBL:X60708; NID:G35335; PIDN:R184001, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.
 J. Biol. Chem. 267, 4824-4833, 1992
 A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer cell lines
 A;Reference number: A42408; MUID:92165847; PMID:1347043
 A;Accession: B42408
 A;Molecule type: mRNA
 A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR1>
 A;Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:G181569; PIDN:AA52308.1; PID:R184001
 A;Experimental source: intestine
 A;Note: this sequence corresponds with the author's translation
 A;Accession: A42408
 A;Molecule type: mRNA
 A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR2>
 A;Cross-references: UNIPARC:UPI0000172A28; GB:M80536; NID:G181569
 A;Note: sequence extracted from NCBI backbone (NCBI:83986, NCBI:83988); this sequence is identical to the one in the GenBank database
 R;Gorvel, J.P.; Ferrero, A.; Chambaud, L.; Rigal, A.; Bonicel, J.; Maroux, S.
 Gastroenterology 101, 618-625, 1991
 A;Title: Expression of sucrose-isomaltase and dipeptidyl peptidase IV in human small intestine
 A;Reference number: A61136; MUID:91317403; PMID:1677636
 A;Accession: B61136
 A;Molecule type: protein
 A;Residues: 1-15, 'X', 17-22 <GOR>
 A;Cross-references: UNIPARC:UPI0000172A2C
 R;Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.
 Biochem. J. 311, 835-843, 1995
 A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a 5' to 3' direction
 A;Reference number: S59510; MUID:96067599; PMID:7487939
 A;Accession: S59510
 A;Molecule type: DNA
 A;Residues: 1-31 <BOB>
 A;Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:G1195574; PIDN:AA35614.1; PID:R184001
 R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, H.
 J. Immunol. 149, 481-486, 1992
 A;Title: Cloning and functional expression of the T cell activation antigen CD26.
 A;Reference number: I56154; MUID:92325476; PMID:1352530
 A;Accession: I56154
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-436, 'S', 438-766 <TAN>
 A;Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:G180082; PIDN:AA51943.1; PID:R184001
 R;Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.
 Immunogenetics 40, 331-338, 1994
 A;Title: Genomic organization, exact localization, and tissue expression of the human CD26 gene
 A;Reference number: S59857; MUID:95012454; PMID:7927537
 A;Accession: S59857
 A;Molecule type: DNA
 A;Residues: 1-436, 'S', 438-766 <ABB>
 A;Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734
 C;Genetics:
 A;Gene: GDB:DP4
 A;Gene-references: GDB:125239; OMIM:102720
 A;Map position: 2q24.3-2q24.3
 A;Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
 C;Superfamily: dipeptidyl-peptidase IV
 C;Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; homodimer; proteinase; transmembrane protein
 F;1-6/Domain: intracellular #status predicted <INT>
 F;7-28/Domain: transmembrane #status predicted <TM>
 F;29-766/Domain: extracellular #status predicted <EXT>
 F;85, 92, 150, 219, 229, 281, 321, 520, 685/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;630, 708, 740/Active site: Ser, Asp, His #status predicted

Query Match 81.5%; Score 53; DB 1; Length 766;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPNG 9
 DB 213 ALWWSPNG 220

RESULT 3

A;Accession: S24313
 dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
 N;Alternate names: GP110; membrane glycoprotein 110K; OX-61
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
 C;Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
 R;Hong, W.; Doyle, D.
 Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
 A;Title: CDNA cloning for a bile canalliculus domain-specific membrane glycoprotein of rat liver
 A;Reference number: A39914; MUID:88068516; PMID:3479775
 A;Accession: A39914
 A;Molecule type: mRNA
 A;Residues: 1-792 <HON>
 A;Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:R184001
 R;Ogata, S.; Misumi, Y.; Ikehara, Y.
 J. Biol. Chem. 264, 3596-3601, 1989
 A;Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA
 A;Reference number: A33315; MUID:89123496; PMID:2563382
 A;Accession: A33315
 A;Molecule type: mRNA
 A;Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623, 'A'
 A;Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41096.1; PID:R184001
 A;Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
 A;Accession: A33315
 A;Molecule type: protein
 A;Residues: 1-20; 35-54; 427-443; 505-509; 511-520; 530-538; 593-600; 602-608; 618-627 <HO2>
 A;Cross-references: UNIPARC:UPI0000172A2E; UNIPARC:UPI0000172A30; UNIPARC:UPI0000172A31
 A36; UNIPARC:UPI0000172A37
 R;McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
 Hepatology 11, 534-544, 1990
 A;Title: Identification of the bile canallicular cell surface molecule GP110 as the ecto-enzyme
 A;Reference number: A60730; MUID:90228896; PMID:1970322
 A;Accession: A60730
 A;Molecule type: protein
 A;Residues: 28-47, 'XX', 50-53, 55-58 <MCC>
 A;Cross-references: UNIPARC:UPI0000172A38
 R;Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
 Biochemistry 31, 2582-2587, 1992
 A;Title: Identification of the active site residues in dipeptidyl peptidase IV by affinity labeling
 A;Reference number: A42203; MUID:92190188; PMID:1347701
 A;Accession: A42203
 A;Molecule type: protein
 A;Residues: 'R', 625-630, 'X', 632-648 <OG2>
 A;Cross-references: UNIPARC:UPI0000172A39
 R;Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
 Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
 A;Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl peptidase
 A;Reference number: S38949; MUID:94128239; PMID:7905271
 A;Accession: S38949
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 281-302 <IWA>
 A;Cross-references: UNIPARC:UPI0000172A3A
 R;Hong, W.; Doyle, D.
 J. Biol. Chem. 263, 16892-16898, 1988
 A;Title: Membrane orientation of rat gp110 as studied by in vitro translation.
 A;Reference number: A31781; MUID:89034185; PMID:3182821
 A;Accession: A31781
 A;Molecule type: mRNA
 A;Residues: 1-40 <HO3>

A:Cross-references: UNIPARC:UPI0000172A3B
 C:Comment: This protein is localized to the bile canalliculus, which is the apical domain
 C:Superfamily: dipeptidyl-peptidase IV
 C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; liver; serine proteina
 F:1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>
 F:1-28/Domain: signal sequence #link MATS #status experimental <SIG>
 F:1-6/Domain: intracellular #status predicted <INT>
 F:7-28/Domain: transmembrane #status predicted <TMN>
 F:729-792/Domain: extracellular #status predicted <EXT>
 F:729-34/Domain: propeptide #link MATS #status experimental <PRO>
 F:35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
 F:83-90,148,217,227,319,521,686/Binding site: carbohydrate (Asn) (covalent) #status pred
 F:631/Active site: Ser #status experimental
 F:709,741/Active site: Asp, His #status predicted

Query Match 81.5%; Score 53; DB 1; Length 792;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWNSPNG 9
 |||||
 Db 211 ALWNSPNG 218

RESULT 4
 S23752
 dipeptidyl-peptidase IV (SC 3.4.14.5) alpha chain - mouse
 N:Alternate names: CD26 alpha subunit; THAM alpha subunit
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S23752; A46465; A56030
 R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darnoul, D.; Naquet, P.; Pierres, M.
 J. Biol. Chem. 267, 2200-2208, 1992
 A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
 A:Reference number: S23752; MUID:92129288; PMID:1370813
 A:Accession: S23752
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-760 <MAR>
 A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
 R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
 J. Immunol. 147, 447-454, 1991
 A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
 A:Reference number: A46465; MUID:91302787; PMID:1712807
 A:Accession: A46465
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <VIV>
 A:Cross-references: UNIPARC:UPI0000172A2E
 A:Experimental source: M14.T thymoma cells, Swiss nu/nu
 A>Note: sequence extracted from NCBI backbone (NCBIP:42236)
 R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
 Biochemistry 33, 15204-15214, 1994
 A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
 A:Reference number: A56030; MUID:95092780; PMID:7999781
 A:Accession: A56030
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 746-760 <BER>
 A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
 C:Genetics:
 A:Gene: CD26
 C:Superfamily: dipeptidyl-peptidase IV
 C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
 F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:624,702,734/Active site: Ser, Asp, His #status predicted

Query Match 72.3%; Score 47; DB 1; Length 760;
 Best Local Similarity 100.0%; Pred. No. 9.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWNSPN 8
 |||||

Db 207 ALWNSPN 213

RESULT 5

T19514

hypothetical protein C27C12.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T19514

R:Thomas, K.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19134

A:Accession: T19514

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-829 <WIL>

A:Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DC53; EMBL:Z69883; PIDN:CAA93743.1.1;

A:Experimental source: clone C27C12

C:Genetics:

A:Gene: CESP:C27C12.7

A:Map position: X

A:Introns: 4/3; 51/1; 123/3; 166/3; 188/2; 279/2; 392/3; 501/3; 553/1; 583/1; 606/2; 647

C:Superfamily: dipeptidyl-peptidase IV

Query Match 72.3%; Score 47; DB 2; Length 829;

Best Local Similarity 66.7%; Pred. No. 10;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWNSPNGK 10

|:|||||:

Db 265 AVWNPSGR 273

RESULT 6

D86146

F22L4.11 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: D86146

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86146

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-509 <STO>

A:Cross-references: UNIPROT:Q9LMM4; UNIPARC:UPI00000A1319; GB:AE005172; NID:98920593; PI

C:Genetics:

A:Map position: 1

C:Superfamily: Arabidopsis hypothetical protein F13M22.23

Query Match 70.8%; Score 46; DB 2; Length 509;

Best Local Similarity 75.0%; Pred. No. 9.1;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWNPSNGK 10

||| |||:

Db 109 LWNXPNGE 116

RESULT 7

AC2988

3-demethylubiquinone-9 3-methyltransferase [imported] - Agrobacterium tumefaciens (strai

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AC2988
 R;Wood, D.W.; Sebubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AC2988
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-250 <KUR>
 A;Cross-references: UNIPROT:Q8UA66; UNIPARC:UPI000013799C; GB:AE008689; PIDN:AAL44321.1;
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: ubiG
 A;Map position: linear chromosome
 C;Superfamily: 3-demethylubiquinone-9 3-O-methyltransferase; bioC homology

Query Match 69.2%; Score 45; DB 2; Length 250;
 Best Local Similarity 77.8%; Pred. No. 6.7;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWWSPTGK 10
 |||||
 Db 22 AEWWSPTGK 30
 |||||

RESULT 8
 E98295
 dihydroxypolyphenylbenzoate methyltransferase (AC004669) [imported] - Agrobacterium tum
 C;Species: Agrobacterium tumefaciens
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C;Accession: E98295
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Accession: E98295
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-250 <KUR>
 A;Cross-references: UNIPROT:Q8UA66; UNIPARC:UPI000013799C; GB:AE007870; PIDN:AAK89887.1;
 C;Genetics:
 A;Gene: AGR_L 2640
 A;Map position: linear chromosome
 C;Superfamily: 3-demethylubiquinone-9 3-O-methyltransferase; bioC homology

Query Match 69.2%; Score 45; DB 2; Length 250;
 Best Local Similarity 77.8%; Pred. No. 6.7;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWWSPTGK 10
 |||||
 Db 22 AEWWSPTGK 30
 |||||

RESULT 9
 T26259
 hypothetical protein W07A8.1 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T26259
 R;Basham, V.
 submitted to the EMBL Data Library, November 1996
 A;Reference number: Z20184
 A;Accession: T26259
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-320 <WIL>
 A;Cross-references: UNIPROT:Q9XUK6; UNIPARC:UPI0000074E93; EMBL:Z82075; PIDN:CAB04930.1;

A;Experimental source: clone W07A8

C;Genetics:

A;Gene: CESP:W07A8.1

A;Map position: 5

A;Introns: 83/1; 141/3; 256/3

C;Superfamily: Caenorhabditis elegans hypothetical protein W07A8.4

Query Match 69.2%; Score 45; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 8.4;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWSPPNG 9
 |||||

Db 74 WWSPPNG 79
 |||||

RESULT 10

T37700

probable dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe) (fragmen

C;Species: Schizosaccharomyces pombe

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000

C;Accession: T37700

R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A;Reference number: Z21737

A;Accession: T37700

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-743 <DEV>

A;Cross-references: UNIPARC:UPI000006A67C; EMBL:Z98596; PIDN:CAB11208.1; GSPDB:GN000066;

A;Experimental source: strain 972h-; cosmid c14C4

C;Genetics:

A;Gene: SPDB:SPAC14C4.15C

A;Map position: 1

Query Match 69.2%; Score 45; DB 2; Length 743;

Best Local Similarity 75.0%; Pred. No. 19;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||

Db 178 AVWWSFDG 185
 |||||

RESULT 11

H64118

4-alpha-glucanotransferase homolog - Haemophilus influenzae (strain Rd KW20)

C;Species: Haemophilus influenzae

C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C;Accession: H64118

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A;Reference number: A64000; MUID:95350630; PMID:7542800

A;Accession: H64118

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-699 <TIGR>

A;Cross-references: UNIPROT:P45176; UNIPARC:UPI000012EB5B; GB:U32815; GB:L42023; NID:91

C;Genetics:

A;Start codon: GTG

C;Superfamily: 4-alpha-glucanotransferase

Query Match 67.7%; Score 44; DB 2; Length 699;

Best Local Similarity 60.0%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10
 :|||

Db 466 FRLLWIPKGGK 475
 :|||

RESULT 12
G75316
probable lipase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: G75316
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75316
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-237 <WHI>
A:Cross-references: UNIPROT:Q9RGP6; UNIPARC:UPI00000D3F31; GB:AE002044; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2078
A:Map position: 1

Query Match 66.2%; Score 43; DB 2; Length 237;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSP 7
||| |||
DB 186 YATWSP 192

RESULT 13
I48673
matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I48673
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce
A:Reference number: I38046; MUID:95224014; PMID:7708715
A:Accession: I48673
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <RNA>
A:Cross-references: UNIPARC:UPI0000030971; EMBL:X83536; NID:g804999; PIDN:CAAS5820.1; PI
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-97/Domain: activation peptide #status predicted <PRO>
F:61-284/Domain: matrix metalloproteinase homology <MMP>
F:313-508/Domain: hemopexin repeat homology <PXN>
F:93-239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status
F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted
F:240/Active site: Glu #status predicted

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
||| |||
DB 418 ALFWMPNGK 426

RESULT 14
I38028
matrix metalloproteinase 14 (EC 3.4.24.-) membrane type precursor - human
N:Alternate names: matrix metalloproteinase MMP14; membrane type matrix metalloproteinase
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004

C:Accession: I38028; G02274; I38046; S78011; S45341; S71384
R:Will, H.; Hinzmann, B.
Eur. J. Biochem. 231, 602-608, 1995
A:Title: CDNA sequence and mRNA tissue distribution of a novel human matrix metalloprotei
A:Reference number: I38028; MUID:95377289; PMID:7649159
A:Accession: I38028
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <WILL>
A:Cross-references: UNIPROT:P50281; UNIPARC:UPI0000048136; EMBL:248481; NID:g963053; PID
R:Luo, G.
submitted to the EMBL Data Library, November 1995
A:Reference number: H00963
A:Accession: G02274
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <LUO>
A:Cross-references: UNIPARC:UPI0000048136; EMBL:U41078; NID:g1127836; PIDN:AAA83770.1; P
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce
A:Reference number: I38046; MUID:95224014; PMID:7708715
A:Accession: I38046
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-7,'S',9-582 <OKA>
A:Cross-references: UNIPARC:UPI000003096F; EMBL:X83535; NID:g804993; PIDN:CAAS58519.1; P
R:Seiki, M.
submitted to the EMBL Data Library, January 1994
A:Reference number: S78011
A:Accession: S78011
A:Molecule type: mRNA
A:Residues: 1-7,'S',9-337,'K',339-582 <SEI>
A:Cross-references: UNIPARC:UPI000003DC76; EMBL:D26512; NID:g793762; PIDN:BAA05519.1; PI
R:Sato, H.; Takino, T.; Okada, Y.; Cao, J.; Shinagawa, A.; Yamamoto, E.; Seiki, M.
Nature 370, 61-65, 1994
A:Title: A matrix metalloproteinase expressed on the surface of invasive tumour cells.
A:Reference number: S45341; MUID:94286011; PMID:8015608
A:Accession: S45341
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-7,'S',9-188,'R',190,'A',192-267,'K',269-272,'HY',275,'P',277-285,'KQ',288,
A:Cross-references: UNIPARC:UPI0000157586; EMBL:D26512
R:Sato, H.; Kinoshita, T.; Takino, T.; Nakayama, K.; Seiki, M.
FEBS Lett. 393, 101-104, 1996
A:Title: Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP)
A:Reference number: S71384; MUID:96397540; PMID:8804434
A:Accession: S71384
A:Molecule type: protein
A:Residues: 112-116 <SAW>
A:Cross-references: UNIPARC:UPI0000175D8E
C:Genetics:
A:Gene: GDB:MMP14; MT1-MMP
A:Cross-references: GDB:375731; OMIM:600754
A:Map position: 14q11-14q12
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei
C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-97/Domain: activation peptide #status predicted <PRO>
F:61-284/Domain: matrix metalloproteinase homology <MMP>
F:98-582/Product: matrix metalloproteinase 14 membrane type #status predicted <MAT>
F:285-313/Domain: hinge #status predicted <HNG>
F:314-508/Domain: hemopexin repeat homology <PXN>
F:539-562/Domain: transmembrane #status predicted <TM>
F:93,239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status
F:130/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted
F:240/Active site: Glu #status predicted
F:319-508/Disulfide bonds: #status predicted

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 15

I84471
matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat
N:Alternate names: membrane type metalloproteinase
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I84471; I61946
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, F.
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce
A:Reference number: I38046; MUID:95224014; PMID:7708715
A:Accession: I84471
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <RES>
A:Cross-references: UNIPROT:Q10739; UNIPARC:UPI0000030970; EMBL:X83537; NID:g805012; PID
A:Accession: I61946
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67, 'M', 69-254, 'A', 256-582 <RE2>
A:Cross-references: UNIPARC:UPI00001679D1; EMBL:X91785; NID:g1001926; PIDN:CAA62897.1; F
C:Genetics:
A:Gene: mt-mmp
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotein
F:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-97/Domain: activation peptide #status predicted <PRO>
F:61-284/Domain: matrix metalloproteinase homology <MMP>
F:313-508/Domain: hemopexin repeat homology <PXN>
F:93,239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status
F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted
F:240/Active site: Glu #status predicted

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 16

AG3275
hexaprenyldihydroxybenzoate methyltransferase (EC 2.1.1.114) [imported] - Brucella melit
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AG3275
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
M.; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3275
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <KUR>
A:Cross-references: UNIPROT:Q8VJ98; UNIPARC:UPI0000057BA9; GB:AE008917; PIDN:AAU51370.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0188
A:Map position: I
C:Superfamily: 3-demethylubiquinone-9-3-O-methyltransferase; bioc homology
C:Keywords: methyltransferase

Query Match 64.6%; Score 42; DB 2; Length 248;
Best Local Similarity 66.7%; Pred. No. 19;

QY 2 ALWWSPNGK 10
||:|||||
Db 418 ALFWMPNGK 426

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 ALWWSPNGK 10
||:|||||
Db 22 AQWMDPOGK 30

RESULT 17

A72050
GlpT/PgpT/UhpT family protein CP0082 [imported] - Chlamydomophila pneumoniae (strains CML0
N:Alternate names: hexosephosphate transport protein
C:Species: Chlamydomophila pneumoniae; Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: A72050; F81616
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: A72050
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <ARN>
A:Cross-references: UNIPROT:Q9Z7N9; UNIPARC:UPI0000137B1F; GB:AE001363; NID:
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81616
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <REA>
A:Cross-references: UNIPARC:UPI0000137B1F; GB:AE002171; GB:AE002161; NID:g7189009; PIDN
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: uhpC; CP0082
C:Superfamily: hexose phosphate transport protein uhpT

Query Match 64.6%; Score 42; DB 2; Length 455;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWWSPNG 9
:|||||
Db 124 FALWMLNG 132

RESULT 18

F86573
hexosephosphate transport [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae; Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86573
R:Shitai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: F86573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <STO>
A:Cross-references: UNIPROT:Q9Z7N9; UNIPARC:UPI0000137B1F; GB:BA000008; NID:g8979037;
A:Experimental source: strain J138
C:Genetics:
A:Gene: uhpC
C:Superfamily: hexose phosphate transport protein uhpT

Query Match 64.6%; Score 42; DB 2; Length 455;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWWSPNG 9

Db 124 FALWGLNG 132 :|||: ||

RESULT 19
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YHR028C
C:Species: Saccharomyces cerevisiae
C:Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C:Accession: S46780; A30107
R:Du, Z.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of *S. cerevisiae* cosmid 8082.
A:Reference number: S46773
A:Accession: S46780
A:Molecule type: DNA
A:Residues: 1-818 <DUZ>
A:Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:G500689; PID
R:Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A:Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in
A:Reference number: A30107; MUID:89174971; PMID:2647766
A:Accession: A30107
A:Molecule type: DNA
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKRGKERKF', 376-57
A:Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A:Note: the authors translated the codon ACC for residue 572 as Asn
C:Genetics:
A:Gene: SGD:DAP2; STE13; MIPS:YHR028C
A:Cross-references: SGD:S0001070; MIPS:YHR028C
A:Map position: 8R
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F:30-45/Domain: transmembrane #status predicted <TM>
F:63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.6%; Score 42; DB 1; Length 818;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWWSPTG 257
Db 250 ALWWSPTG 257
:|||||

RESULT 20
T01868
Probable glucose-6-phosphate/phosphate-translocator T24W8.5 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01868
R:Latreille, P.; Elliott, G.; Le, T.
submitted to the EMBL Data Library, August 1998
A:Description: The sequence of *A. thaliana* T24W8.
A:Reference number: Z14449
A:Accession: T01868
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-246 <LAT>
A:Cross-references: UNIPROT:O81514; UNIPARC:UPI00000A73B0; EMBL:AF077409; NID:G3319365;
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 17/3; 65/3; 80/1; 95/1; 170/2; 188/2
A:Note: T24W8.5
C:Superfamily: triose phosphate/3-phosphoglycerate/phosphate translocator

Query Match 63.1%; Score 41; DB 2; Length 246;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSPTG 9

Db 22 FAIWWALNG 30 :|||: ||

RESULT 21
E87353
3-demethylubiquinone-9 3-methyltransferase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87353
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <STO>
A:Cross-references: UNIPROT:Q9A9X1; UNIPARC:UPI000013799D; GB:AE005673; NID:G13422093; P
C:Genetics:
A:Gene: CC0840
C:Superfamily: 3-demethylubiquinone-9 3-O-methyltransferase; bioC homology

Query Match 63.1%; Score 41; DB 2; Length 252;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALWWSPTG 10
Db 26 AEWDPK GK 34
:|||||

RESULT 22
I38029
matrix metalloproteinase 15 (EC 3.4.24.-) membrane type precursor - human
N:Alternate names: membrane-type matrix metalloproteinase 2 (MT2MMP); MMP15
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
C:Accession: I38029; G00041
R:Will, H.; Hinzmann, B.
Eur. J. Biochem. 231, 602-608, 1995
A:Title: cDNA sequence and mRNA tissue distribution of a novel human matrix metalloprotei
A:Reference number: I38028; MUID:95377289; PMID:7649159
A:Accession: I38029
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-669 <WIL>
A:Cross-references: UNIPROT:P51511; UNIPARC:UPI000003DC75; EMBL:Z48482; NID:G963055; PID
R:Seiki, M.
submitted to GenBank, July 1996
A:Reference number: H00041
A:Accession: G00041
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 106-608, 'R', 610-669 <SEI>
A:Cross-references: UNIPARC:UPI00000705A7; GB:D86331; NID:G1418214; PIDN:BAA13071.1; PID
C:Genetics:
A:Gene: GDB:MMP15; MT2-MMP
A:Cross-references: GDB:642246
C:Superfamily: interstitial collagenase; metalloproteinase; zinc; zymogen
C:Keywords: glycoprotein; hydrolase; metalloproteinase; signal sequence #status predicted <SIG>
F:1-41/Domain: signal sequence #status predicted <ACT>
F:42-125/Domain: activation peptide #status predicted <ACT>
F:79-304/Domain: matrix metalloproteinase homology <MMP>
F:133-669/Product: matrix metalloproteinase 15 #status predicted <MAT>
F:364-559/Domain: hemopexin repeat homology <XPN>
F:624-644/Domain: transmembrane #status predicted <TRM>
F:111,259,263,269/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status
F:150/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:259,263,269/Binding site: zinc, catalytic (His) (active) #status predicted
F:260/Active site: Glu #status predicted

Mon Mar 13 10:54:37 2006

```
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A>Title: Differential expression of two distinct forms of mRNA encoding members of a di
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:G408719; PIDN
A>Note: sequence extracted from NCBI backbone (NCBI:75138)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein
F:257,342/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.1%; Score 41; DB 2; Length 803;
Best Local Similarity 75.0%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
DB 241 AHWWSPPG 248

RESULT 26
I54331
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I54331
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A>Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-rela
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I54331
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-865 <RES>
A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:G306705; PIDN
C:Superfamily: dipeptidyl-peptidase IV

Query Match 63.1%; Score 41; DB 2; Length 865;
Best Local Similarity 75.0%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
DB 303 AHWWSPPG 310

RESULT 27
JQ2166
spindle body protein precursor - Heliothis armigera poxvirus
C:Species: Heliothis armigera poxvirus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: JQ2166
R:Dall, D.; Srikantha, A.; Vera, A.; Lai-Fook, J.; Symonds, T.
J. Gen. Virol. 74, 1811-1818, 1993
A>Title: A gene encoding a highly expressed spindle body protein of Heliothis armigera
A:Reference number: JQ2166; MUID:93389435; PMID:8376960
A:Accession: JQ2166
A:Molecule type: DNA
A:Residues: 1-351 <DAL>
A:Cross-references: UNIPROT:Q05894; UNIPARC:UPI0000135E3A; GB:L08077; NID:G329565; PIDN
C:Superfamily: spheroidin
C:Keywords: glycoprotein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-351/Product: spindle body protein #status predicted <MAT>
F:195/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.5%; Score 40; DB 2; Length 351;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WWSPPNG 9
DB 4 WWSPPNG 9
```

```
Query Match 63.1%; Score 41; DB 2; Length 669;
Best Local Similarity 62.5%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
DB 469 AIWWEPTG 476

RESULT 23
T01541
hypothetical protein A_IG005110.16 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01541
R:Andrews, S.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of A. thaliana IG005110.
A:Reference number: Z14347
A:Accession: T01541
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-785 <AND>
A:Cross-references: UNIPROT:O23072; UNIPARC:UPI00000A0893; EMBL:AF013293; NID:G2252823;
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 284/3; 467/3
A>Note: A_IG005110.16

Query Match 63.1%; Score 41; DB 2; Length 785;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WWSPPNG 9
DB 104 WKPNG 109

RESULT 24
I68600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I68600
R:Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A>Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-rela
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PIDN:AAA35761.1; PID:
C:Superfamily: dipeptidyl-peptidase IV

Query Match 63.1%; Score 41; DB 2; Length 803;
Best Local Similarity 75.0%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
DB 241 AHWWSPPG 248

RESULT 25
A41793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimazaki, S.
```

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T30040
R;Gattung, S.; Wu, X.

submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid F20A1.

A;Reference number: Z20726

A;Accession: T30040

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-466 <GAT>

A;Cross-references: UNIPROT:Q19614; UNIPARC:UPI000017B912; EMBL:U53150; PIDN:AAA96130.1;

A;Experimental source: strain Bristol N2; clone F20A1

C;Genetics:

A;Gene: CESP:F20A1.9

A;Map position: 5

A;Introns: 45/3; 122/1; 271/3; 294/2; 315/3; 381/3; 423/2

Query Match 61.5%; Score 40; DB 2; Length 466;

Best Local Similarity 62.5%; Pred. No. 73;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LWSPNGK 10

Db 166 IMWSKNGE 173

RESULT 31

AF0998

phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49) - Salmonella enterica subsp. enter

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Oct-2003

C;Accession: AF0998

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Cross-references: UNIPARC:UPI000005A7B8; GB:AL513382; PIDN:CAD08114.1; PID:gi6505093;

C;Genetics:

A;Gene: STV4296

A;Superfamily: phosphoenolpyruvate carboxykinase [ATP]

C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 61.5%; Score 40; DB 2; Length 539;

Best Local Similarity 75.0%; Pred. No. 83;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LWSPNGK 10

Db 81 LWSDKKG 88

RESULT 32

T20450

hypothetical protein E04D5.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T20450

R;McMurray, A.

submitted to the EMBL Data Library, October 1995

A;Reference number: Z19277

A;Accession: T20450

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-618 <WIL>

A;Cross-references: UNIPROT:Q19052; UNIPARC:UPI0000076B22; EMBL:Z66496; PIDN:CAA91279.1;

Db 41 WWPNG 46

RESULT 28

JC5185

enhancing factor precursor - armyworm (Pseudaletia separata)

C;Species: Pseudaletia separata

C;Date: 16-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000

A;Accession: JC5185; PC4254

R;Hayakawa, T.; Xu, J.; Hukuvara, T.

Gene 177, 269-270, 1996

A;Title: Cloning and sequencing of the gene for an enhancing factor from Pseudaletia sep

A;Reference number: JC5185; MUID:97080535; PMID:8921880

A;Accession: JC5185

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-352 <HAY1>

A;Cross-references: UNIPARC:UPI00000F4377; DBDJ:D50590; NID:GI741865; PIDN:BAA09138.1; E

A;Accession: PC4254

A;Status: preliminary

A;Molecule type: protein

A;Residues: 21-45;183-201 <HAY2>

A;Cross-references: UNIPARC:UPI00001784AE; UNIPARC:UPI00001784AF

C;Genetics:

A;Gene: ef

C;Superfamily: spheroidin

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-352/Product: enhancing factor #status predicted <MAT>

Query Match 61.5%; Score 40; DB 2; Length 352;

Best Local Similarity 83.3%; Pred. No. 56;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WWSPNG 9

Db 41 WWPNG 46

RESULT 29

T18825

hypothetical protein C01G10.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C;Accession: T18825

R;Matthews, L.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19027

A;Accession: T18825

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-421 <WIL>

A;Cross-references: UNIPARC:UPI0000082BD8; EMBL:Z81030; PIDN:CAB02716.1; GSPDB:GN000023;

A;Experimental source: clone C01G10

C;Genetics:

A;Gene: CESP:C01G10.10

A;Map position: 5

A;Introns: 31/2; 72/3; 122/2; 177/1; 272/3; 315/1; 353/2

C;Superfamily: O-sialoglycoprotein endopeptidase

Query Match 61.5%; Score 40; DB 2; Length 421;

Best Local Similarity 71.4%; Pred. No. 66;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSN 8

Db 356 ALWWSN 362

RESULT 30

T30040

hypothetical protein F20A1.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

A;Experimental source: clone E04D5
C;Genetics:
A;Gene: CESP:E04D5.1
A;Map position: 2
A;Introns: 57/1; 81/2; 107/2; 151/3; 212/1; 305/3; 586/3
C;Superfamily: translation initiation factor eIF-2A

Query Match 61.5%; Score 40; DB 2; Length 618;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
DB 328 YAAKWPNGR 337
|||:||||

RESULT 33
G70928
probable 4-ALPHA-GLUCANOTRANSFERASE - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: G70928
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Scars, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70928
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-724 <COL>
A;Cross-references: UNIPROT:O53932; UNIPARC:UPI000012EB5C; GB:AL022021; GB:AL123456; NID
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV1781C
C;Superfamily: 4-alpha-glucanotransferase

Query Match 61.5%; Score 40; DB 2; Length 724;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
DB 482 FRUWIPDG 490
|||:||||

RESULT 34
F90109
splicing factor Prp8 [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: F90109
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Red
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: F90109
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2057 <DOU>
A;Cross-references: UNIPROT:Q9AW36; UNIPARC:UPI000000A2A43; GB:AJ010592; NID:g12580717; F
C;Genetics:
A;Map position: 2
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 61.5%; Score 40; DB 2; Length 2057;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSP 7
DB 1351 FTLWSP 1357
:|||||

RESULT 35
S44625
C50C3.6 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S44625
R;Favella, A.D.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid C50C3.
A;Reference number: S44627
A;Accession: S44625
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2329 <FAV>
A;Cross-references: UNIPROT:P34369; UNIPARC:UPI000013B834; EMBL:L14433; NID:g289649; PFI
C;Genetics:
A;Introns: 277/2; 356/3; 430/1; 678/2; 1150/3; 1484/3; 1642/2; 2191/3; 2311/3
C;Keywords: nucleus

Query Match 61.5%; Score 40; DB 2; Length 2329;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSP 7
DB 1527 FTLWSP 1533
:|||||

RESULT 36
T06077
splicing factor PRP8 homolog T9A14.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06077
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.;
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15184
A;Accession: T06077
A;Molecule type: DNA
A;Residues: 1-2352 <BEV>
A;Cross-references: UNIPROT:O9T016; UNIPARC:UPI00000A4642; EMBL:AL035656; GSPDB:GN00062
A;Experimental source: cultivar Columbia; BAC clone T9A14
C;Genetics:
A;Gene: ATSP:T9A14.60
A;Map position: 4
A;Introns: 94/3; 132/3; 158/3; 289/1; 453/2; 717/1; 806/3; 911/3; 1013/2; 1165/1; 1207/1

Query Match 61.5%; Score 40; DB 2; Length 2352;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSP 7
DB 1553 FTLWSP 1559
:|||||

RESULT 37
B96832
hypothetical protein F18B13.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B96832
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.E.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B96832
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2359 <STO>
A:Cross-references: UNIPROT:Q9SSD2; UNIPARC:UPI000009D7F0; GB:AE005173; NID:g5902365; PI
C:Genetics:
A:Gene: F18B13.15
A:Map position: 1

Query Match 61.5%; Score 40; DB 2; Length 2359;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSWP 7
: |||||
Db 1558 FTLWWSWP 1564

RESULT 38
T38841
probable pre-mRNA splicing factor - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A:Accession: T38841
R:Genies, S.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z21815
A:Accession: T38841
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2363 <GEN>
A:Cross-references: UNIPROT:O14187; UNIPARC:UPI000006A413; EMBL:Z98530; PIDN:CAB11062.1;
A:Experimental source: strain 972h; cosmid c4f8
C:Genetics:
A:Gene: SPDB:SPAC4F8.12c
A:Map position: 1

Query Match 61.5%; Score 40; DB 2; Length 2363;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSWP 7
: |||||
Db 1558 FTLWWSWP 1564

RESULT 39
S34670
splicing factor PRP8 - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: DBF3 protein; protein YHR165c
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
A:Accession: S34670; S48905; S55355; S53613; S51873
R:Hodges, P.E.; Jackson, S.P.; Brown, J.D.; Beggs, J.D.
submitted to the EMBL Data Library, July 1993
A:Description: Predicted protein sequence of the yeast splicing factor PRP8 and of a Nem
A:Reference number: S34670
A:Accession: S34670
A:Molecule type: DNA
A:Residues: 1-2413 <HOD>
A:Cross-references: UNIPROT:P33334; UNIPARC:UPI000013213D; EMBL:Z24732; NID:g395156; PID
R:Macri, C.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9986.
A:Reference number: S46673
A:Accession: S48905
A:Molecule type: DNA
A:Residues: 1-2413 <MAC>

A:Cross-references: UNIPARC:UPI000013213D; EMBL:U00027; NID:g551319; PID:g551320; MIPS:Y
R:Hodges, P.E.; Jackson, S.P.; Brown, J.D.; Beggs, J.D.
Yeast 11, 337-342, 1995
A:Title: Extraordinary sequence conservation of the PRP8 splicing factor.
A:Reference number: S55355; MUID:95304847; PMID:7785334
A:Accession: S55355
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-2413 <HOW>
A:Cross-references: UNIPARC:UPI000013213D; EMBL:Z24732; NID:g395156; PIDN:CAA80854.1; PI
R:Shea, J.E.; Toyn, J.H.; Johnston, L.H.
Nucleic Acids Res. 22, 5555-5564, 1994
A:Title: The budding yeast US snRNP Prp8 is a highly conserved protein which links RNA s
A:Reference number: S53613; MUID:95140615; PMID:7838707
A:Accession: S53613
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-387, 'LYIYIPGPVQCAYHGIIICRVLSRTMSTRRL', 421-1131, 'S', 1133-1574, 'C', 1576-24
A:Cross-references: UNIPARC:UPI0000168BCL; EMBL:L29421; NID:g460041; PIDN:AAA67044.1; PI
C:Genetics:
A:Gene: SGD:PRP8; DBF3
A:Cross-references: SGD:S0001208; MIPS:YHR165c
A:Map position: 8R
C:Function:
A:Description: pre-mRNA splicing
C:Keywords: nucleus; pre-mRNA splicing

Query Match 61.5%; Score 40; DB 2; Length 2413;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSWP 7
: |||||
Db 1606 FTLWWSWP 1612

RESULT 40
T35629
hypothetical protein SC6G9.33 - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
A:Accession: T35629
R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T35629
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-168 <SE>
A:Cross-references: UNIPROT:Q9YAE9; UNIPARC:UPI00000DB18F; EMBL:AL079356; PIDN:CAB45624.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB:SC6G9.33

Query Match 60.0%; Score 39; DB 2; Length 168;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWWSWPNG 9
: |||||
Db 110 ALWWSWPAG 117

RESULT 41
G87558
conserved hypothetical protein CC2496 [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
A:Accession: G87558
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87558

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-185 <STO>

A:Cross-references: UNIPROT:Q9A5F4; UNIPARC:UPI00000C7750; GB:AE005673; NID:gl3424051; F

C:Genetics:

A:Gene: CC2496

Query Match 60.0%; Score 39; DB 2; Length 185;

Best Local Similarity 75.0%; Pred. No. 43;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWSPNG 9

|||||

Db 131 ALWVPAG 138

RESULT 42

T19589

hypothetical protein C30H6.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T19589

R:Morimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19148

A:Accession: T19589

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-223 <WIL>

A:Cross-references: UNIPROT:Q45276; UNIPARC:UPI000008256C; EMBL:Z81044; PIDN:CAB02810.1;

A:Experimental source: clone C30H6

C:Genetics:

A:Gene: CESP:C30H6.1

A:Map position: 4

A:Introns: 28/1; 153/1; 171/3

Query Match 60.0%; Score 39; DB 2; Length 223;

Best Local Similarity 83.3%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWSPN 8

|||||

Db 188 LWQPN 193

RESULT 43

B84282

8-oxoguanine DNA glycosylase [imported] - *Halobacterium* sp. NRC-1

C:Species: *Halobacterium* sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004

C:Accession: B84282

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leitchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B84282

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-301 <STO>

A:Cross-references: UNIPROT:Q9HQ94; UNIPARC:UPI0000063847; GB:AE004437; NID:g10580792; F

C:Genetics:

A:Gene: O9g

C:Superfamily: N-glycosylase/DNA lyase

Query Match 60.0%; Score 39; DB 2; Length 301;

Best Local Similarity 71.4%; Pred. No. 69;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWSPNG 9

|||||

Db 28 LWRPDG 34

RESULT 44

F64610

hypothetical protein HP0726 - *Helicobacter pylori* (strain 26695)

C:Species: *Helicobacter pylori*

C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004

C:Accession: F64610

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: F64610

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-305 <TOM>

A:Cross-references: UNIPROT:O25426; UNIPARC:UPI00000C088D; GB:AE000585; GB:AE000511; NII

C:Superfamily: *Helicobacter* hypothetical protein HP0726

Query Match 60.0%; Score 39; DB 2; Length 305;

Best Local Similarity 75.0%; Pred. No. 70;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSPN 8

|||||

Db 213 YAILWSPN 220

RESULT 45

T08544

hypothetical protein F27B13.70 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 05-Oct-2004

C:Accession: T08544

R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Ma;

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16442

A:Accession: T08544

A:Molecule type: DNA

A:Residues: 1-321 <BEV>

A:Cross-references: UNIPROT:Q9SZQ5; UNIPARC:UPI00000A3578; EMBL:AL050352; GSPDB:GN00062

A:Experimental source: cultivar Columbia; BAC clone F27B13

C:Genetics:

A:Gene: ATSP:F27B13.70

A:Map position: 4

A:Introns: 113/3

Query Match 60.0%; Score 39; DB 2; Length 321;

Best Local Similarity 100.0%; Pred. No. 73;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WSPNGK 10

|||||

Db 166 WSPNGK 171

RESULT 46

S22366

phenylalanine-tRNA ligase (EC 6.1.1.20) alpha chain [validated] - *Thermus aquaticus*

N:Alternate names: phenylalanyl-tRNA synthetase alpha chain

C:Species: *Thermus aquaticus*

C>Date: 22-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 03-Jun-2002

C:Accession: S22366; S25171; S61093; T52502; S21468

R:Keller, B.; Kast, P.; Hennecke, H.

FEBS Lett. 301, 83-88, 1992

A:Title: Cloning and sequence analysis of the phenylalanyl-tRNA synthetase genes (phest

A;Reference number: S22366; MUID:93083630; PMID:1451792
A;Accession: S22366
A;Molecule type: DNA
A;Residues: 1-350 <REL>
A;Cross-references: UNIPARC:UPI00001107A3; EMBL:Z12118; NID:G48253; PIDN:CAA78104.1; PIDN:CAA78104.1; PIDN:CAA78104.1
A;Note: the source is designated as *Thermus thermophilus*
R;Kreutzer, R.; Kruft, V.; Bobkova, E.V.; Lavrik, O.I.; Sprinzl, M.
Nucleic Acids Res. 20, 4173-4178, 1992
A;Title: Structure of the phenylalanyl-tRNA synthetase genes from *Thermus thermophilus* H
A;Reference number: S2517; MUID:92375722; PMID:1508711
A;Accession: S2517
A;Molecule type: DNA
A;Residues: 1-350 <REL>
A;Cross-references: UNIPARC:UPI00001107A3; EMBL:X65609; NID:G48250; PIDN:CAA46559.1; PIDN:CAA46559.1; PIDN:CAA46559.1
A;Experimental source: strain HB8 (ATCC 27634)
A;Note: the source is designated as *Thermus thermophilus*
A;Accession: S61093
A;Molecule type: protein
A;Residues: 1-29 <KRW>
A;Cross-references: UNIPARC:UPI0000176201
R;Lechler, A.; Kreutzer, R.
J. Mol. Biol. 278, 897-901, 1998
A;Title: The phenylalanyl-tRNA synthetase specifically binds DNA.
A;Reference number: Z26096
A;Accession: T52502
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-350 <LEC>
A;Cross-references: UNIPARC:UPI00001107A3; EMBL:Y15464; PIDN:CAA75644.1
A;Experimental source: strain HB8
C;Genetics:
A;Gene: phs
C;Function:
A;Description: EC 6.1.1.20 [validated, MUID:98263257]; binds specifically certain DNA se
A;Superfamily: phenylalanine-tRNA ligase alpha chain
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 60.0%; Score 39; DB 2; Length 350;
Best Local Similarity 50.0%; Pred. NO. 79;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VALWSPNGK 10
:|:|:|:|
Db 267 FAVWPEGK 276

RESULT 47
T40673
homoserine dehydrogenase - fission yeast (*Schizosaccharomyces pombe*)
C;Species: *Schizosaccharomyces pombe*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Dec-2004
C;Accession: T40673
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Wedler, H.; Kutzner, M.; Wambutt, R.
submitted to the EMBL Data Library, January 1999
A;Reference number: Z21944
A;Accession: T40673
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-376 <LYN>
A;Cross-references: UNIPROT:Q94671; UNIPARC:UPI0000129356; EMBL:AL035263; PIDN:CAA22876.
A;Experimental source: strain 972h-; cosmid c776
C;Genetics:
A;Gene: SPDB:SPBC776.03
A;Map position: 2
A;Introns: 14/1
C;Superfamily: homoserine dehydrogenase lacking ACT domain; homoserine dehydrogenase hom

Query Match 60.0%; Score 39; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. NO. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WSPNGK 10
:|:|:|:|

Db 191 WSPNGK 196

RESULT 48

SYZJAL

5-aminolevulinate synthase (EC 2.3.1.37) - *Bradyrhizobium japonicum*

N;Alternate names: delta-aminolevulinic synthase

C;Species: *Bradyrhizobium japonicum*

C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004

C;Accession: A27478

R;McClung, C.R.; Somerville, J.E.; Guerinot, M.L.; Chelam, B.K.

Gene 54, 133-139, 1987

A;Title: Structure of the *Bradyrhizobium japonicum* gene hema encoding 5-aminolevulinic a

A;Reference number: A27478; MUID:87277426; PMID:3609750

A;Accession: A27478

A;Molecule type: DNA

A;Residues: 1-409 <MCC>

A;Cross-references: UNIPROT:P08262; UNIPARC:UPI000016E709; GB:M16751; NID:G152096; PIDN:

C;Comment: This pyridoxal phosphate enzyme catalyzes the formation of 5-aminolevulinic a

C;Genetics:

A;Gene: hema

C;Superfamily: 5-aminolevulinic synthase; glycine C-acetyltransferase homology

C;Keywords: acyltransferase; coenzyme A; phosphoprotein; porphyrin biosynthesis; pyridox

F:50-385/Domain: glycine C-acetyltransferase homology <GCA>

F:247/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

F;389-424/Domain: WD repeat homology <WD4>
F;433-465/Domain: WD repeat homology <WD5>

Query Match 60.0%; Score 39; DB 2; Length 465;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
:|||||
Db 353 YTVWSPNG 361

RESULT 50

T40172
Glucose inhibited division protein a - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40172
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, August 1998
A;Reference number: 221909
A;Accession: T40172
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-666 <LYN>
A;Cross-references: UNIPROT:Q13670; UNIPARC:UPI0000013920F; EMBL:AL031262; PIDN:CAA20319.
A;Experimental source: strain 972h; cosmid c30B4
C;Genetics:
A;Gene: SPDB:SPBC3084.06c
A;Map position: 2
C;Superfamily: gida protein

Query Match 60.0%; Score 39; DB 2; Length 666;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WSPNG 9
:|||||
Db 324 WYTPNG 329

RESULT 51

S61694
fluculation suppression protein SFL1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O3339; protein YOR140w; protein YOR3339w
C;Species: Saccharomyces cerevisiae
C;Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: S61694; JQ0384; S67025
R;Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Sander, C.; Valencia
submitted to the EMBL Data Library, December 1995

A;Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
A;Reference number: S61694
A;Accession: S61694
A;Molecule type: DNA
A;Residues: 1-766 <BSN>
A;Cross-references: UNIPROT:P20134; UNIPARC:UPI00001358AF; EMBL:X94335; NID:g1262139; PI
R;Fujioka, A.; Kikuchi, Y.; Kuhara, S.; Misumi, Y.; Matsumoto, S.; Kobayashi, H.
Gene 85, 321-328, 1989
A;Title: Domains of the SFL1 protein of yeasts are homologous to Myc oncoproteins or yea
A;Reference number: JQ0384; MUID:90185205; PMID:2697640
A;Accession: JQ0384
A;Molecule type: DNA
A;Residues: 1-445; FVQYQFSQ, 455-460, 'KQ', 463-766 <FUJ>
A;Cross-references: UNIPARC:UPI000017459F
R;Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansorge, W.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66965

A;Accession: S67025
A;Molecule type: DNA
A;Residues: 1-766 <VOS>
A;Cross-references: UNIPARC:UPI00001358AF; EMBL:Z75047; NID:g1420352; PIDN:CAA99338.1; F
A;Experimental source: strain S288C
C;Comment: This protein is involved in cell surface assembly and regulation of the gene

C;Genetics:

A;Gene: SGD:SFL1; MIPS:YOR140w
A;Cross-references: MIPS:YOR140w; SGD:S0005666
A;Map position: 15R

C;Superfamily: fluculation suppression protein SFL1; HSF DNA-binding domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;65-170/Domain: HSF DNA-binding domain homology <HSF>

Query Match 60.0%; Score 39; DB 1; Length 766;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LWWSPNG 9
:|||||
Db 82 IWWTPSG 88

RESULT 52

C86427
hypothetical protein F12P21.10 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86427
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C86427
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-802 <STO>
A;Cross-references: UNIPROT:Q9C751; UNIPARC:UPI00000AA76F; GB:AE005172; NID:g1120775;
C;Genetics:
A;Map position: 1

Query Match 60.0%; Score 39; DB 2; Length 802;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WSPNG 8
:|||||
Db 515 WWSNG 519

RESULT 53

T49796
probable sepB protein [imported] - Neurospora crassa

N;Alternate names: protein B9J10.330
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 05-Oct-2004
C;Accession: T49796
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022

A;Accession: T49796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-879 <SCH>
A;Cross-references: UNIPROT:Q9P549; UNIPARC:UPI000006C5C3; EMBL:ALJ56324; GSPDB:GN00116
A;Experimental source: BAC clone B9J10; strain OR74A
C;Genetics:
A;Gene: NCSP:B9J10.330
A;Map position: 6
A;Introns: 619/3

Query Match 60.0%; Score 39; DB 2; Length 879;

```
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALMWSPNG 9
   |||||
Db 239 ALSWSPNG 246

RESULT 54
T33236
carboxypeptidase C (EC 3.4.16.5) precursor [validated] - fission yeast (Schizosaccharomyces
N:Alternate names: carboxypeptidase Y
C:Species: Schizosaccharomyces pombe
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 05-Oct-2004
C:Accession: T33236; T37997
R:Tabuchi, M.; Iwaihara, O.; Ohtani, Y.; Ohuchi, N.; Sakurai, J.; Morita, T.; Iwahara, S
J. Bacteriol. 179, 4179-4189, 1997
A:Title: Vacuolar protein sorting in fission yeast: cloning, biosynthesis, transport, an
A:Reference number: 223357; MUID:97352672; PMID:9209031
A:Accession: T33236
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1002 <TAB>
A:Cross-references: UNIPROT:O13849; UNIPARC:UPI00001271DE; EMBL:D86560; NID:G3046860; PI
R:Oliver, K.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z21760
A:Accession: T37997
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1002 <OLI>
A:Cross-references: UNIPARC:UPI00001271DE; EMBL:Z97209; PIDN:CAB10121.1; GSPDB:GN00066;
A:Experimental source: strain 972h(-); cosmid c19C12
C:Comment: The mature carboxypeptidase exists as a heterodimer held together by a disulf
C:Genes:
A:Gene: cpy1; SPAC19G12.10C
A:Map position: 1
C:Superfamily: Serine carboxypeptidase
C:Keywords: disulfide bond; glycoprotein; heterodimer; hydrolase; serine carboxypeptidase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-521/Domain: propeptide #status predicted <PRO>
F:200-332/Region: 13-residue repeats (H-H-[EK]-P-G-E-H-M-P-P-P-[MF])
F:344-425/Region: 9-residue repeats (E-H-H-[RKQ]-G-G-[KE]-[DE]-K)
F:522-1002/Product: carboxypeptidase C #status experimental <MAT>
F:627-880/Disulfide bonds: #status predicted
F:659/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:715,921,978/Active site: Ser, Asp, His #status predicted

Query Match 60.0%; Score 39; DB 2; Length 1002;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WSPNGK 10
   |||||
Db 952 WSPNGK 957

RESULT 55
AD1280
SNF2-type helicase homolog lmo1644 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1280
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1280
A:Status: preliminary
```

```
A:Molecule type: DNA
A:Residues: 1-1072 <GLA>
A:Cross-references: UNIPROT:Q8Y6P0; UNIPARC:UPI000005500C; GB:NC_003210; PIDN:CAC99722.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1644

Query Match 60.0%; Score 39; DB 2; Length 1072;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YALWWSWP 7
   |||||
Db 989 YDLWNP 995

RESULT 56
AD1643
SNF2-type helicase homolog lmo1685 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1643
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1643
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <GLA>
A:Cross-references: UNIPROT:Q92B65; UNIPARC:UPI000000CC622; GB:AL592022; PIDN:CAC96916.1
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lmo1685

Query Match 60.0%; Score 39; DB 2; Length 1072;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YALWWSWP 7
   |||||
Db 989 YDLWNP 995

RESULT 57
E88451
protein K10D2.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E88451
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E88451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1223 <STO>
A:Cross-references: UNIPROT:Q09589; UNIPARC:UPI000007D4F0; GB:chr_III; PIDN:AAA62539.1;
C:Genetics:
A:Gene: K10D2.1
A:Map position: 3

Query Match 60.0%; Score 39; DB 2; Length 1223;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 5 WSPNGK 10
 DB 429 WSPNGK 434

RESULT 58

T30875
 PRP8 protein homolog - Trypanosoma brucei
 C:Species: Trypanosoma brucei
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T30875
 RiDuecke, S.; Kloeckner, T.; Palfi, Z.; Boshart, M.; Bindereif, A.
 EMBO J. 16, 4433-4440, 1997
 A:Title: Trans mRNA splicing in trypanosomes: cloning and analysis of a PRP8-homologous
 A:Reference number: Z20915; MUID:97392473; PMID:9250687
 A:Accession: T30875
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2403 <LUE>
 A:Cross-references: UNIPROT:O15881; UNIPARC:UPI0000080703; EMBL:Y12638; NID:e1044898; PI
 C:Genetics:
 A:Note: p277

Query Match 60.0%; Score 39; DB 2; Length 2403;
 Best Local Similarity 71.4%; Pred. NO. 4.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSP 7
 DB 1595 FALWMP 1601

RESULT 59

T18378
 variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N:Alternate names: erythrocyte membrane binding protein 1 (EMPI)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18378
 RiBaruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.
 Cell 82, 77-87, 1995
 A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
 A:Reference number: Z18925; MUID:95330812; PMID:7541722
 A:Accession: T18378
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2924 <BAR>
 A:Cross-references: UNIPROT:Q25733; UNIPARC:UPI0000079077; EMBL:U27338; NID:g914918; PI
 C:Genetics:
 A:Gene: EMP1
 A:Introns: 2476/3

Query Match 60.0%; Score 39; DB 2; Length 2924;
 Best Local Similarity 71.4%; Pred. NO. 5.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WSPNGK 10
 DB 1096 WNWNGK 1102

RESULT 60

A46112
 genome polyprotein - rice tungro spherical virus (strain Los Banos)
 C:Species: rice tungro spherical virus
 C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
 C:Accession: A46112
 RiShen, P.; Kaniewska, M.; Smith, C.; Beachy, R.N.
 Virology 193, 621-630, 1993
 A:Title: Nucleotide sequence and genomic organization of rice tungro spherical virus.
 A:Reference number: A46112; MUID:93212494; PMID:8460478
 A:Accession: A46112
 A:Molecule type: Genomic RNA

A:Residues: 1-3473 <SHE>
 A:Cross-references: UNIPROT:Q83034; UNIPARC:UPI00000F231A; GB:S57835
 C:Superfamily: rice tungro spherical virus genome polyprotein
 C:Keywords: glycoprotein; polyprotein
 F:106,367,903,1018,1274,1311,1385,1665,1755,2291,2779/Binding site: carbohydrate (Asn)

Query Match 60.0%; Score 39; DB 1; Length 3473;
 Best Local Similarity 62.5%; Pred. NO. 7e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWSPNGK 10
 DB 1145 IWYDNGK 1152

Search completed: March 11, 2006, 12:25:43
 Job time : 24.7895 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:12:17 ; Search time 97.8947 Seconds
(without alignments)
72.070 Million cell updates/sec

Title: US-10-774-242A-4
Perfect score: 65
Sequence: 1 VALWSPNGK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	760	1	SEPR HUMAN
2	65	100.0	760	2	Q53TP5 HUMAN
3	60	92.3	350	2	Q6P7D6 RAT
4	60	92.3	755	2	Q91651 XENLA
5	60	92.3	761	1	SEPR MOUSE
6	60	92.3	761	2	Q8R492 RAT
7	55	84.6	751	2	Q75982 AGKHB
8	55	84.6	751	2	Q75983 AGKHB
9	54	83.1	639	2	Q68K44 XHEPC
10	53	81.5	759	2	Q52181 CHICK
11	53	81.5	765	2	Q5R7G7 PONPY
12	53	81.5	766	1	DPP4 HUMAN
13	53	81.5	766	1	DPP4 PIG
14	53	81.5	766	2	Q53TN1 HUMAN
15	53	81.5	767	1	DPP4 RAT
16	53	81.5	817	2	Q6FPV7 CANGA
17	53	81.5	883	2	Q55H13 CRYNE
18	53	81.5	883	2	Q5K721 CRYNE
19	52	80.0	872	2	Q75C44 ASHGO
20	51	78.5	439	2	Q6PG25 XENLA
21	49	75.4	404	2	Q82C69 STRAW
22	49	75.4	737	2	Q641D6 XENLA
23	49	75.4	748	2	P70092 XENLA
24	49	75.4	868	2	Q6CFX2 YARLI
25	48	73.8	249	1	UBIG RHILLO
26	48	73.8	704	2	Q4S309 TETNG
27	48	73.8	794	2	Q60J70 CAEBR
28	47	72.3	553	2	Q4SKD2 TETNG
29	47	72.3	760	1	DPP4 MOUSE
30	47	72.3	765	1	DPP4 BOVIN
31	47	72.3	765	1	DPP4 FELCA
					Q9n217 f dipeptid

32	47	72.3	779	2	Q4HVM6 GIBZE	Q4hvm6 gibberella
33	47	72.3	829	2	Q6CNM1_KJULA	Q6cnm1 kluyveromyc
34	47	72.3	829	2	Q18253 CAEBL	Q18253 caenorhabdi
35	47	72.3	836	2	Q61PC7 CAEBR	Q61pc7 caenorhabdi
36	46	70.8	240	2	Q4FNA2_9RICK	Q4fna2 candidatus
37	46	70.8	509	2	Q9LMM4_ARATH	Q9lmm4 arabidopsis
38	46	70.8	771	2	Q42812_ASPOR	Q42812 aspergillus
39	46	69.2	217	2	Q7UKM8_RHOBA	Q7ukm8 rhodopirell
40	45	69.2	320	1	UBIG_AGR75	Q8ua66 agrobacteri
41	45	69.2	320	2	Q9XUK6 CAEBL	Q9xuk6 caenorhabdi
42	45	69.2	579	2	Q7T290_BRARE	Q7t290 brachydanio
43	45	69.2	787	2	Q4LY82_9BURK	Q4ly82 burkholderi
44	45	69.2	853	1	YDZF_SCHPO	Q9p7e9 schizosacch
45	44	67.7	227	2	Q6FWQ2_XENLA	Q6fwq4 xenopus lae
46	44	67.7	322	2	Q9ZN63_PPRU	Q9zn63 prevotella
47	44	67.7	383	2	Q7UP27_RHOBA	Q7up27 rhodopirell
48	44	67.7	476	2	Q4Q7R7_LEIMA	Q4q7r7 leishmania
49	44	67.7	575	2	Q5USC3_XENLA	Q5usc3 xenopus lae
50	44	67.7	576	2	Q6DCW8_XENLA	Q6dcw8 xenopus lae
51	44	67.7	578	2	Q5BL52_XENTR	Q5bl52 xenopus tro
52	44	67.7	699	1	MALQ_HAEI8	P45176 haemophilus
53	44	67.7	699	2	Q4QK66_HAEI8	Q4qk66 haemophilus
54	44	67.7	1256	2	Q82L10_STRAW	Q82l10 streptomyce
55	44	67.7	1271	2	Q7QT29_GIALA	Q7qt29 giardia lam
56	44	67.7	2320	2	Q9XYM9_TRIVA	Q9xym9 trichomonas
57	44	67.7	2427	2	Q4FW74_LEIMA	Q4fw74 leishmania
58	43	66.2	148	2	Q97657_CANFA	Q97657 canis famil
59	43	66.2	148	2	Q6QLX0_CANFA	Q6qlx0 canis famil
60	43	66.2	222	1	LFTR_LEGPA	Q544e4 legionella
61	43	66.2	222	1	LFTR_LEGPH	Q5zum5 legionella
62	43	66.2	222	1	LFTR_LEGPL	Q5wv89 legionella
63	43	66.2	234	1	LFTR_ERWCT	Q6d3u2 erwilia car
64	43	66.2	234	2	Q4K3V9_PSEF5	Q4k3v9 pseudomonas
65	43	66.2	236	1	LFTR_IDILO	Q5r0b5 idiomarina
66	43	66.2	237	2	Q9RSP6_DEIRA	Q9rsp6 deinococcus
67	43	66.2	241	2	Q6SHR2_9BACT	Q6shr2 uncultured
68	43	66.2	247	1	UBIG_BARHE	Q6gsk3 bartonella
69	43	66.2	249	1	LFTR_XANCP	Q8p996 xanthomonas
70	43	66.2	249	2	Q4UUK2_XANCP	Q4uuk2 xanthomonas
71	43	66.2	508	2	Q4ZU12_PSESY	Q4zu12 pseudomonas
72	43	66.2	560	2	Q4J5H2_AZOFI	Q4j5h2 azotobacter
73	43	66.2	580	1	MMP14_PIG	Q9xt90 sus scrofa
74	43	66.2	582	1	MMP14_HUMAN	P50281 homo sapien
75	43	66.2	582	1	MMP14_MOUSE	P53690 mus musculu
76	43	66.2	582	1	MMP14_RABIT	Q95220 oryctolagus
77	43	66.2	582	1	MMP14_RAT	Q10739 rattus norv
78	43	66.2	582	2	Q6GSF3_HUMAN	Q6gsf3 homo sapien
79	43	66.2	582	2	Q9GLE4_BOVIN	Q9gle4 bos taurus
80	43	66.2	582	2	Q5RES1_PONPY	Q5res1 pongo pygma
81	43	66.2	582	2	Q8BTX2_MOUSE	Q8btx2 mus musculu
82	43	66.2	582	2	Q6DFU5_MOUSE	Q6dfu5 mus musculu
83	43	66.2	582	2	Q6IN06_RAT	Q6in06 rattus norv
84	43	66.2	582	2	Q9LV35_ARATH	Q9lv35 arabidopsis
85	43	66.2	609	2	Q4RR35_TETNG	Q4rr35 tetraodon n
86	43	66.2	715	2	Q6BP08_DEBHA	Q6bp08 debaryomyce
87	43	66.2	870	2	Q59SM7_CANAL	Q59sm7 candida alb
88	43	66.2	211	2	Q56335_9PICO	Q56335 foot-and-mo
89	42	64.6	142	2	Q56330_9PICO	Q56330 foot-and-mo
90	42	64.6	144	2	Q56336_9PICO	Q56336 foot-and-mo
91	42	64.6	144	2	Q56333_9PICO	Q56333 foot-and-mo
92	42	64.6	146	2	Q56344_9PICO	Q56344 foot-and-mo
93	42	64.6	146	2	Q9WLR3_9PICO	Q9wlr3 foot-and-mo
94	42	64.6	146	2	Q56334_9PICO	Q56334 foot-and-mo
95	42	64.6	147	2	Q56337_9PICO	Q56337 foot-and-mo
96	42	64.6	147	2	Q56340_9PICO	Q56340 foot-and-mo
97	42	64.6	147	2	Q56341_9PICO	Q56341 foot-and-mo
98	42	64.6	147	2	Q56342_9PICO	Q56342 foot-and-mo
99	42	64.6	147	2	Q56345_9PICO	Q56345 foot-and-mo
100	42	64.6	147	2	Q92839_9PICO	Q92839 foot-and-mo
101	42	64.6	147	2	Q92840_9PICO	Q92840 foot-and-mo
102	42	64.6	147	2	Q92841_9PICO	Q92841 foot-and-mo
103	42	64.6	147	2	Q9WLR2_9PICO	Q9wlr2 foot-and-mo
104	42	64.6	147	2	Q9WLR2_9PICO	Q9wlr2 foot-and-mo

105	42	64.6	148	2	O56343_9PICO	O56343	foot-and-mo	178	41	63.1	570	2	O75MF0_HUMAN	O75mf0	homo sapien
106	42	64.6	149	2	O56329_9PICO	O56329	foot-and-mo	179	41	63.1	571	2	O4HGP3_CAMCO	O4hgp3	campylobact
107	42	64.6	149	2	O56332_9PICO	O56332	foot-and-mo	180	41	63.1	573	2	O560D7_CRYNE	O560d7	cryptococcu
108	42	64.6	149	2	O56332_9PICO	O56332	foot-and-mo	181	41	63.1	573	2	O5KPE0_CRYNE	O5kpe0	cryptococcu
109	42	64.6	149	2	O56338_9PICO	O56338	foot-and-mo	182	41	63.1	577	2	O7PNC7_ANOGA	O7pnc7	anopheles g
110	42	64.6	149	2	O56339_9PICO	O56339	foot-and-mo	183	41	63.1	617	2	O5ISN3_MACFA	O5isn3	macaca faec
111	42	64.6	216	2	O80JY0_9PICO	O80jy0	foot-and-mo	184	41	63.1	657	1	NMP15_MOUSE	O54732	mus musculus
112	42	64.6	216	2	O80JY0_9PICO	O80jy6	foot-and-mo	185	41	63.1	658	1	O76LU8_ORYLA	O76lu8	oryzias lat
113	42	64.6	216	2	O80JY0_9PICO	O80jy7	foot-and-mo	186	41	63.1	669	1	NMP15_HUMAN	P51511	homo sapien
114	42	64.6	216	2	O80JY8_9PICO	O80jy8	foot-and-mo	187	41	63.1	682	1	O7X717_ORISA	O7x717	oryza sativ
115	42	64.6	216	2	O80JZ3_9PICO	O80jz3	foot-and-mo	188	41	63.1	711	2	O9LY16_STRCO	O9lly6	streptomyce
116	42	64.6	216	2	O80JZ4_9PICO	O80jz4	foot-and-mo	189	41	63.1	724	2	O5IS95_SAIMIRI	O5is95	saimiri bol
117	42	64.6	238	2	O9M9A4_ARATH	O9m9a4	arabidopsis	190	41	63.1	740	2	O82CF9_STRAW	O82cf9	streptomyce
118	42	64.6	241	2	O6SHD4_9BACT	O6shd4	uncultured	191	41	63.1	765	2	O4WPH9_ASFPU	O4wph9	aspergillus
119	42	64.6	248	1	UBIG_BARQU	O69011	bartonella	192	41	63.1	765	2	O14425_ASFPU	O14425	aspergillus
120	42	64.6	248	1	UBIG_BRUME	O89011	brucella	193	41	63.1	785	2	O23072_ARATH	O23072	arabidopsis
121	42	64.6	248	1	UBIG_BRUSE	O89011	brucella me	194	41	63.1	801	2	O5U4C2_MOUSE	O5u4c2	mus musculus
122	42	64.6	248	1	O5LWM6_SILPO	O5lwm6	silicibacte	195	41	63.1	803	1	DPP6_PANTR	O51850	pan troglod
123	42	64.6	276	2	O7XQM0_ORYSA	O7xqm0	oryza sativ	196	41	63.1	803	2	O9QV78_WMURI	O9qvt8	rattus sp.
124	42	64.6	360	2	O69112_92ZZZ	O69i12	uncultured	197	41	63.1	804	1	DPP6_MOUSE	O80vms	mus musculus
125	42	64.6	360	2	O69126_92ZZZ	O69i26	uncultured	198	41	63.1	815	2	O4RK63_TETNG	O4rk63	tetraodon n
126	42	64.6	360	2	O69140_92ZZZ	O69i40	uncultured	199	41	63.1	842	2	O4RVF8_TETNG	O4rvf8	tetraodon n
127	42	64.6	371	2	O69130_92ZZZ	O69i30	uncultured	200	41	63.1	859	1	DPP6_RAT	P46101	rattus norv
128	42	64.6	372	2	O69132_92ZZZ	O69i32	uncultured	201	41	63.1	863	1	DPP6_BOVIN	P42659	bos taurus
129	42	64.6	374	2	O69118_92ZZZ	O69i18	uncultured	202	41	63.1	863	1	DPP6_HUMAN	P42658	homo sapien
130	42	64.6	455	1	UHPT_CHLPN	O927n9	chlamydia p	203	41	63.1	865	1	O8CC14_MOUSE	O8cc14	mus musculus
131	42	64.6	455	2	O5L737_CHLAB	O5l737	chlamydogphi	204	41	63.1	1050	2	O5R9P6_PONPY	O5r9p6	pongo pygma
132	42	64.6	455	2	O824R5_CHLCV	O824r5	chlamydogphi	205	41	63.1	1089	2	O8BQX1_MOUSE	O8bqx1	mus musculus
133	42	64.6	517	2	O7SFT7_NEUCR	O7sft7	neurospora	206	41	63.1	1129	2	O8IMI9_HUMAN	O8imi9	homo sapien
134	42	64.6	604	2	O53S84_HUMAN	O53s84	homo sapien	207	41	63.1	1139	2	O61YH3_CAEBR	O61yh3	caenorhabdi
135	42	64.6	618	2	O4SF93_TETNG	O4sf93	tetraodon n	208	41	63.1	1201	2	O4S1P2_TETNG	O4s1p2	tetraodon n
136	42	64.6	712	1	IF39_ARATH	O96521	arabidopsis	209	41	63.1	1286	2	O4SVZ2_TETNG	O4svz2	tetraodon n
137	42	64.6	714	2	O8GUM1_ARATH	O8gum1	arabidopsis	210	41	63.1	1803	2	O4SUR5_TETNG	O4sur5	tetraodon n
138	42	64.6	717	2	O6C1H8_YARLI	O6c1h8	yarrowia li	211	41	63.1	2025	2	O6A056_MOUSE	O6a056	mus musculus
139	42	64.6	775	2	O5J6J3_TRIRU	O5j6j3	trichophyto	212	41	63.1	2309	2	O9QXJ5_MOUSE	O9qxj5	mus musculus
140	42	64.6	792	2	O8XXL7_RALSO	O8xxl7	raletonia s	213	41	63.1	3006	2	O6QF20_MYCO	O6qf20	mycobacteri
141	42	64.6	796	1	DPP10_HUMAN	O8n608	homo sapien	214	40	61.5	211	2	SOST_MOUSE	O59p68	mus musculus
142	42	64.6	796	1	DPP10_RAT	O6q629	rattus norv	215	40	61.5	211	2	O5YCW2_LACPL	O5ycw2	lactobacill
143	42	64.6	796	2	O6INB7_XENLA	O6inb7	xenopus lae	216	40	61.5	213	1	SOST_RAT	O89p67	rattus norv
144	42	64.6	797	1	DPP10_MOUSE	O6nxb7	mus musculus	217	40	61.5	220	2	O8S5V8_ORYSA	O8s5v8	oryza sativ
145	42	64.6	801	2	O4RK66_TETNG	O4rk66	tetraodon n	218	40	61.5	288	2	O52852_BACPU	O52852	bacillus pu
146	42	64.6	809	2	O4NVV1_9DELT	O4nvv1	anaeromyxob	219	40	61.5	346	2	O4XY20_PLACH	O4xy20	plasmodium
147	42	64.6	818	1	DAP2_YEAST	P19862	saccharomyc	220	40	61.5	351	1	SPIN_HABPV	O05894	heliothis a
148	42	64.6	818	2	O6SR87_YEAST	O6sr87	saccharomyc	221	40	61.5	352	2	O12686_PPOXV	O12686	pseudaletia
149	42	64.6	860	2	O4RQJ4_TETNG	O4rqj4	tetraodon n	222	40	61.5	406	2	O8H7S5_ORYSA	O8h7s5	oryza sativ
150	42	64.6	930	2	O59M48_CANAL	O59m48	candida alb	223	40	61.5	428	2	O4RPX5_TETNG	O4rp5	caenorhabdi
151	42	64.6	1199	2	O4JXY9_CORJK	O4jxy9	corynebacte	224	40	61.5	431	2	O94CQ0_ORYSA	O94cq0	oryza sativ
152	42	64.6	2335	2	O719N0_9PICO	O719n0	foot-and-mo	225	40	61.5	433	2	O19614_CAEEL	O19614	caenorhabdi
153	41	63.1	121	2	O6Q225_9BACT	O6q225	uncultured	226	40	61.5	436	2	O6JH02_HUMAN	O6jh02	homo sapien
154	41	63.1	126	2	O9XSI0_BOVIN	O9xsi8	bos taurus	227	40	61.5	437	2	O6JH03_HUMAN	O6jh03	homo sapien
155	41	63.1	136	2	O80IG1_9PICO	O80ig1	foot-and-mo	228	40	61.5	458	2	O6NEJ2_CORDI	O6nej2	corynebacte
156	41	63.1	136	2	O80IG2_9PICO	O80ig2	foot-and-mo	229	40	61.5	463	2	O4QU36_9HIV1	O4qu36	human immun
157	41	63.1	136	2	O80IG3_9PICO	O80ig3	foot-and-mo	230	40	61.5	463	2	O4QU34_9HIV1	O4qu34	human immun
158	41	63.1	136	2	O80IG4_9PICO	O80ig4	foot-and-mo	231	40	61.5	463	2	O4QU32_9HIV1	O4qu32	human immun
159	41	63.1	136	2	O80IG5_9PICO	O80ig5	foot-and-mo	232	40	61.5	471	2	O7WUW4_BORPE	O7wu4	bordecella
160	41	63.1	149	2	O56328_9PICO	O56328	foot-and-mo	233	40	61.5	471	2	O7W3Z7_BORPA	O7w3z7	bordecella
161	41	63.1	155	2	O72UJ6_LEPIC	O72uj6	leptospira	234	40	61.5	471	2	O7WFC9_BORBA	O7wfc9	bordecella
162	41	63.1	155	2	O8F0F7_LEPIN	O8f0f7	leptospira	235	40	61.5	471	2	O7WFC9_BORBA	O7wfc9	bordecella
163	41	63.1	165	2	O7UXP7_RHOBA	O7uxp7	rhodopirell	236	40	61.5	475	2	O84DL6_BORNOE	O84dl6	oenococcu
164	41	63.1	216	2	O80JY9_9PICO	O80jy9	foot-and-mo	237	40	61.5	482	1	BGBP2_MANSE	O61sb6	manduca sex
165	41	63.1	231	2	O7Q8W1_ANOGA	O7q8w1	anopheles g	238	40	61.5	482	1	O6LRD4_PHOPR	O6lrd4	photobacter
166	41	63.1	236	1	UBIG_HAEDU	O7qvk2	haemophilus	239	40	61.5	484	2	O6LRD4_PHOPR	P41033	salmonella
167	41	63.1	246	2	O81514_ARATH	O81514	arabidopsis	240	40	61.5	538	1	PECK_SALTU	O55v41	cryptococcu
168	41	63.1	248	1	UBIG_RHIME	O92mk1	rhizobium m	241	40	61.5	538	2	O5KL40_CRYNE	O5kl40	cryptococcu
169	41	63.1	252	1	UBIG_CAUCR	O9a9x1	caulobacter	242	40	61.5	539	1	PECK_SALTU	O57ix4	salmonella
170	41	63.1	252	2	O4WEJ4_ASFPU	O4wej4	aspergillus	243	40	61.5	539	2	O5PLX8_SALPA	O5plx8	salmonella
171	41	63.1	252	2	O5BEG5_SCHJA	O5db95	schistosoma	244	40	61.5	540	2	O5E1X3_VIBF1	O5e1x3	vibrio fisc
172	41	63.1	253	1	UBIG_BRAJA	O89xuz	bradyrhizob	245	40	61.5					
173	41	63.1	449	2	O7SF86_NEUCR	O7sf86	neurospora	246	40	61.5					
174	41	63.1	458	2	O5A3W6_CANAL	O5a3w6	candida alb	247	40	61.5					
175	41	63.1	462	2	O8SD29_9CAUD	O8sd29	pseudomonas	248	40	61.5					
176	41	63.1	479	2	O8FE81_ECCL6	O8fe81	escherichia	249	40	61.5					
177	41	63.1	564	2	O7KZY0_HUMAN	O7kzy0	homo sapien	250	40	61.5					

251	40	61.5	547	2	Q55W2_CRYNE	Q55W22 cryptococcus	324	39	60.0	238	2	Q8F5W8_COREF	Q8faw8 corynebacte
252	40	61.5	548	2	Q7UJW7_RHOBA	Q7ujw7 rhodopirell	325	39	60.0	253	2	Q7N0K8_PHOLL	Q7n0k8 photorhabdu
253	40	61.5	566	2	Q5KJ66_CRYNE	Q5kje6 cryptococcus	326	39	60.0	275	2	Q73Z74_MYCPA	Q73z74 mycobacteri
254	40	61.5	570	2	Q19052_CAEEL	Q19052 caenorhabdi	327	39	60.0	301	2	Q8HQ94_HALSA	Q8hq94 halobacteri
255	40	61.5	597	2	Q61ZHI_CAEER	Q61zhi caenorhabdi	328	39	60.0	303	2	P92932_ALISA	P92932 allium sati
256	40	61.5	623	2	Q6UAQ0_TENG	Q6uaq0 tetraodon n	329	39	60.0	305	2	Q25426_HELPY	Q25426 helicobacte
257	40	61.5	640	2	Q8AXA5_ACASC	Q8axa5 acanthopagr	330	39	60.0	307	2	Q7RZE8_NEUCR	Q7rze8 neurospora
258	40	61.5	647	2	Q71MN2_SPAAU	Q71mn2 sparus aura	331	39	60.0	317	2	Q6A9F1_PROAC	Q6a9f1 propionibac
259	40	61.5	719	1	IF39_TOBAC	P56821 nicotiana t	332	39	60.0	321	2	Q9SZQ5_ATH	Q9szq5 streptomyce
260	40	61.5	724	1	MAUQ_MYCBO	P65337 mycobacteri	333	39	60.0	327	2	Q9RPT6_9ACTO	Q9rpt6 arbidopsi
261	40	61.5	724	1	MAUQ_MYCTU	P65336 mycobacteri	334	39	60.0	338	2	Q6ZFU0_ORYSA	Q6zfu0 oryza sativ
262	40	61.5	727	2	Q6IAL1_CAEER	Q6ial1 caenorhabdi	335	39	60.0	340	2	Q69PT1_ORYSA	Q69pt1 oryza sativ
263	40	61.5	800	2	Q61GQ9_CAEER	Q61gg9 caenorhabdi	336	39	60.0	343	2	Q7S214_NEUCR	Q7s214 neurospora
264	40	61.5	809	2	Q9NAN8_CAEEL	Q9nan8 caenorhabdi	337	39	60.0	344	2	Q8G369_9VIRU	Q8g369 rice tungro
265	40	61.5	829	2	Q51B00_ENTHI	Q51b00 entamoeba h	338	39	60.0	346	2	Q8QRL4_9VIRU	Q8qrl4 rice tungro
266	40	61.5	837	2	Q5AN11_DICDI	Q5an11 dictyosteli	339	39	60.0	346	2	Q55XW9_CRYNE	Q55xw9 cryptococcu
267	40	61.5	911	2	Q4IAY7_GIBZE	Q4iay7 gibberella	340	39	60.0	346	2	Q5KM26_CRYNE	Q5km26 cryptococcu
268	40	61.5	1055	2	Q5KAP5_CRYNE	Q5kap5 cryptococcus	341	39	60.0	350	1	SYFA_THET2	Q72hal thermus the
269	40	61.5	1055	2	Q55KT2_CRYNE	Q55kt2 cryptococcus	342	39	60.0	350	1	SYFA_THET8	Q5sgx2 thermus the
270	40	61.5	1328	2	Q96OV2_DROME	Q96ov2 drosophila	343	39	60.0	350	1	SYFA_THET8	Q72001 thermus the
271	40	61.5	1656	2	Q21948_CAEEL	Q21948 caenorhabdi	344	39	60.0	376	1	Q94671_SCHPO	Q94671 schizosacch
272	40	61.5	1706	2	Q619J3_CAEER	Q619j3 caenorhabdi	345	39	60.0	392	2	Q7W040_BORPE	Q7w040 bordetella
273	40	61.5	2057	2	Q9AW36_GUTH	Q9aw36 guillardia	346	39	60.0	405	2	Q6FED5_ACIAD	Q6fed5 acinetobact
274	40	61.5	2270	2	Q512S8_ENTHI	Q512s8 entamoeba h	347	39	60.0	408	2	Q72VL2_LEPIC	Q72vl2 leptospira
275	40	61.5	2307	2	Q61213_CAEER	Q61213 caenorhabdi	348	39	60.0	408	2	Q8F968_LEPIN	Q8f968 leptospira
276	40	61.5	2317	2	Q962X1_PARTE	Q962x1 paramecium	349	39	60.0	409	1	HEM1_BRAJA	P08462 bradyrhizob
277	40	61.5	2327	2	Q8T295_DICDI	Q8t295 dictyosteli	350	39	60.0	409	1	Q6USX9_KLEPN	Q6usx9 klebsiella
278	40	61.5	2328	2	Q4PEC7_USPWA	Q4pec7 ustilago na	351	39	60.0	410	2	Q7UPL7_RHOBA	Q7upl7 rhodopirell
279	40	61.5	2329	1	PRP8_CAEEL	P34369 caenorhabdi	352	39	60.0	419	1	TOUB_LEGPH	Q5x421 legionella
280	40	61.5	2335	1	PRP8_HUMAN	Q6p2q9 homo sapien	353	39	60.0	419	1	TOUB_LEGPH	Q5zv69 legionella
281	40	61.5	2335	1	PRP8_MOUSE	Q99pv0 mus musculu	354	39	60.0	422	2	Q7W3V1_BORPA	Q7w3v1 bordetella
282	40	61.5	2335	1	Q7ZX28_XENLA	Q7zx28 xenopus lae	355	39	60.0	422	2	Q7WF81_BORBR	Q7wf81 bordetella
283	40	61.5	2341	2	Q5CIV6_CRYHO	Q5ci66 cryptococcus	356	39	60.0	439	2	Q7NWR4_CHRVO	Q7nwr4 chromobacte
284	40	61.5	2350	2	Q5SIR5_ORYSA	Q5sir5 oryza sativ	357	39	60.0	443	1	CORO7_RAT	Q35828 rattus norv
285	40	61.5	2350	2	Q5WAB4_ORYSA	Q5wab4 oryza sativ	358	39	60.0	444	2	Q6ALD5_CIOSA	Q6ald9 ciona savig
286	40	61.5	2352	2	Q9T016_ATH	Q9t016 arbidopsi	359	39	60.0	449	2	Q7V0E3_PROMP	Q7v0e3 prochloroco
287	40	61.5	2356	2	Q6CCC5_YARLI	Q6ccc5 yarrowia li	360	39	60.0	463	2	Q7UKY8_RHOBA	Q7uky8 rhodopirell
288	40	61.5	2359	2	Q9SSD2_ATH	Q9ssd2 arbidopsi	361	39	60.0	465	1	PRP4_YEAST	P20053 saccharomyc
289	40	61.5	2363	2	Q14187_SCHPO	Q14187 schizosacch	362	39	60.0	465	2	Q75FL6_LEPIC	Q75fl6 leptospira
290	40	61.5	2363	2	Q520V6_MAGGR	Q520v6 magnaporthe	363	39	60.0	465	2	Q8EXJ8_LEPIC	Q8exj8 leptospira
291	40	61.5	2370	2	Q4IJS2_GIBZE	Q4ij22 gibberella	364	39	60.0	471	2	Q6ALG3_XENTR	Q6alg3 xenopus tro
292	40	61.5	2374	2	Q7SBD3_NEUCR	Q7sbd3 neurospora	365	39	60.0	479	2	Q6NTZ6_XENLA	Q6ntz6 xenopus lae
293	40	61.5	2379	2	Q5CUG7_CRYPV	Q5cug7 cryptospori	366	39	60.0	494	2	Q6BYC3_DESHA	Q6byc3 debaryomyce
294	40	61.5	2398	2	Q7Q1I3_ANOQA	Q7q1i3 anopheles g	367	39	60.0	515	2	Q4IAR0_GIBZE	Q4iar0 gibberella
295	40	61.5	2396	2	Q9V664_DROME	Q9v664 drosophila	368	39	60.0	515	2	Q6NU48_XENLA	Q6nu48 xenopus lae
296	40	61.5	2402	2	Q75DL5_ASHGO	Q75dl5 ashaya goss	369	39	60.0	529	2	Q6CAI4_YARLI	Q6cai4 yarrowia li
297	40	61.5	2411	2	Q6CXV6_KLUFA	Q6cxv6 kluyveromyc	370	39	60.0	530	2	Q5NHK5_FRATT	Q5nhk5 franciella
298	40	61.5	2413	1	PRP8_YEAST	F33334 saccharomyc	371	39	60.0	553	2	Q4W935_ASPFU	Q4w935 aspergillus
299	40	61.5	2414	2	Q6FOZ4_CANGA	Q6foz4 candida gla	372	39	60.0	556	2	Q52801_ORYSA	Q52801 oryza sativ
300	40	61.5	2416	2	Q5A322_CANAL	Q5a322 candida alb	373	39	60.0	596	2	Q582R9_9TRYP	Q582r9 trypanosoma
301	40	61.5	2433	2	Q6BQU7_DEBHA	Q6bqu7 debaryomyce	374	39	60.0	596	2	Q6K880_ORYSA	Q6k880 oryza sativ
302	40	61.5	2736	2	Q4N062_THEPA	Q4n062 theileria p	375	39	60.0	596	2	Q7AC03_GEOSL	Q7ac03 geobacter s
303	40	61.5	2786	2	Q4UCB6_THEAN	Q4ucb6 theileria a	376	39	60.0	605	2	Q4RNC5_TENG	Q4rnc5 tetraodon n
304	40	61.5	2852	2	Q419P3_GIBZE	Q419p3 gibberella	377	39	60.0	610	2	Q8GUG3_ATH	Q8gug3 arbidopsi
305	40	61.5	2945	2	Q5BAK7_EWENT	Q5bak7 aspergillus	378	39	60.0	656	2	Q5XWR4_LEGPL	Q5xwr4 legionella
306	40	61.5	2958	2	Q7RWZ3_PLAYO	Q7rn23 plasmodium	379	39	60.0	656	2	Q5ZV01_LEGPH	Q5zv01 legionella
307	40	61.5	3136	2	Q81IX5_PLAF7	Q8ilx5 plasmodium	380	39	60.0	666	1	Y079_SCHPO	Q13670 schizosacch
308	40	61.5	3170	2	Q4W155_ASPFU	Q4w155 aspergillus	381	39	60.0	680	2	Q8SSL8_ENCCU	Q8ssl8 enccephalito
309	39.5	60.8	138	2	Q6QF29_9MYCO	Q6qf29 mycobacteri	382	39	60.0	703	2	Q8DU73_9VIRU	Q8du73 rice tungro
310	39.5	60.8	147	2	Q6QEV7_MYCXB	Q6qev7 mycobacteri	383	39	60.0	703	2	Q5W1L9_9VIRU	Q5w1l9 rice tungro
311	39.5	60.8	177	2	Q6QF05_9MYCO	Q6qf05 mycobacteri	384	39	60.0	703	2	Q98650_9VIRU	Q98650 rice tungro
312	39.5	60.8	149	2	Q5S588_MYCXB	Q5s588 mycobacteri	385	39	60.0	703	2	Q98650_9VIRU	Q98650 rice tungro
313	39	60.0	55	2	Q75HU5_ORYSA	Q75hj5 oryza sativ	386	39	60.0	713	2	Q73ZV1_MYCPA	Q73zv1 mycobacteri
314	39	60.0	136	2	Q5QLZ1_ORYSA	Q5qlz1 oryza sativ	387	39	60.0	713	2	Q8S7Q0_ORYSA	Q8s7q0 oryza sativ
315	39	60.0	154	2	Q4R9R8_TETNG	Q4r9r8 tetraodon n	388	39	60.0	719	2	Q9N341_CAEEL	Q9n341 caenorhabdi
316	39	60.0	168	2	Q9XAE9_STRCO	Q9xae9 streptomyce	389	39	60.0	728	2	Q6ZS54_HUMAN	Q6zss4 homo sapien
317	39	60.0	179	2	Q628C2_CAEER	Q628c2 caenorhabdi	390	39	60.0	733	2	Q5FOY6_GLUOX	Q5fgy6 gluconobact
318	39	60.0	185	2	Q9AS44_CAUCR	Q9as44 caulobacter	391	39	60.0	739	2	Q8G6V5_BIFLO	Q8g6v5 bifidobacte
319	39	60.0	197	2	Q5Z289_NOCEA	Q5z289 nocardia fa	392	39	60.0	766	1	SFL1_YEAST	P20134 saccharomyc
320	39	60.0	202	2	Q7XPV9_ORYSA	Q7xpv9 oryza sativ	393	39	60.0	802	2	Q9C751_ATH	Q9c751 arbidopsi
321	39	60.0	205	2	Q6SHG0_9BACT	Q6shg0 uncultured	394	39	60.0	854	2	Q4HYV9_GIBZE	Q4hyv9 gibberella
322	39	60.0	223	2	Q45276_CAEEL	Q45276 caenorhabdi	395	39	60.0	879	2	Q9P549_NEUCR	Q9p549 neurospora
323	39	60.0	234	1	UBIG_COXBU	Q820b5 coxiella bu	396	39	60.0				

397	39	60.0	899	2	QBH6R9	PONTR	QBh6r9	poncirus tr	470	38.5	59.2	147	2	Q6QF02_9MYCO	Q6qf02	mycobacteri
398	39	60.0	922	1	COB07	MOUSE	Q942v7	mus musculus	471	38.5	59.2	147	2	Q6QF03_9MYCO	Q6qf03	mycobacteri
399	39	60.0	948	2	QB8X3	DRBHA	Q69x3	debariomyce	472	38.5	59.2	147	2	Q6QF07_9MYCO	Q6qf07	mycobacteri
400	39	60.0	951	2	Q4PBV9	USTMA	Q6pbv9	ustiliago ma	473	38.5	59.2	147	2	Q6QF11_9MYCO	Q6qf11	mycobacteri
401	39	60.0	957	2	Q1G07	CABEL	Q81g07	caenorhabdi	474	38.5	59.2	147	2	Q6QF12_MYCFO	Q6qf12	mycobacteri
402	39	60.0	1002	1	CBPY	SCHPO	Q13849	schizosacch	475	38.5	59.2	147	2	Q6QF13_MYCFO	Q6qf13	mycobacteri
403	39	60.0	1064	2	P94295	BACCE	P94295	bacillus ce	476	38.5	59.2	147	2	Q6QF15_9MYCFA	Q6qf15	mycobacteri
404	39	60.0	1064	2	Q4MUK6	BACCE	Q4muk6	bacillus ce	477	38.5	59.2	147	2	Q6QF16_MYCFA	Q6qf16	mycobacteri
405	39	60.0	1064	2	Q6HKE1	BACHK	Q6hke1	bacillus th	478	38.5	59.2	147	2	Q6QF19_9MYCO	Q6qf19	mycobacteri
406	39	60.0	1064	2	Q81F57	BACCR	Q81f57	bacillus ce	479	38.5	59.2	147	2	Q6QF22_MYCCH	Q6qf22	mycobacteri
407	39	60.0	1064	2	Q73AA4	BACCR	Q73aa4	bacillus ce	480	38.5	59.2	147	2	Q6QF24_9MYCO	Q6qf24	mycobacteri
408	39	60.0	1064	2	Q81S67	BACAN	Q81s67	bacillus an	481	38.5	59.2	147	2	Q6QF36_9MYCO	Q6qf36	mycobacteri
409	39	60.0	1064	2	Q63D05	BACZ	Q63d05	bacillus ce	482	38.5	59.2	147	2	Q6QF37_9MYCO	Q6qf37	mycobacteri
410	39	60.0	1065	2	Q830T4	ENTFA	Q830t4	enterococu	483	38.5	59.2	147	2	Q6QF38_MYCAG	Q6qf38	mycobacteri
411	39	60.0	1072	2	Q71Z25	LISMF	Q71z25	listeria mo	484	38.5	59.2	147	2	Q6QF40_MYCAB	Q6qf40	mycobacteri
412	39	60.0	1072	2	Q8Y6P0	LISMO	Q8y6p0	listeria mo	485	38.5	59.2	153	2	Q581X8_9MYCO	Q581x8	mycobacteri
413	39	60.0	1072	2	Q92B65	LISIN	Q92b65	listeria in	486	38.5	59.2	156	2	Q581Y3_9MYCO	Q581y3	mycobacteri
414	39	60.0	1085	2	Q811I7	MOUSE	Q811i7	mus musculus	487	38.5	59.2	163	2	Q59518_MYCFO	Q59518	mycobacteri
415	39	60.0	1165	2	Q4T2J8	TETNG	Q4t2j8	tetradodon n	488	38.5	59.2	163	2	Q59562_MYCHA	Q59562	mycobacteri
416	39	60.0	1175	2	Q73HF4	WOLPM	Q73hf4	wolbachia p	489	38.5	59.2	167	2	Q5VHM9_MYCSM	Q5vhm9	mycobacteri
417	39	60.0	1223	2	Q09589	CABEL	Q09589	caenorhabdi	490	38.5	59.2	167	2	Q5VHN0_9MYCO	Q5vhn0	mycobacteri
418	39	60.0	1335	2	Q4SHU4	TETNG	Q4shu4	tetradodon n	491	38.5	59.2	167	2	Q5VHN1_9MYCO	Q5vhn1	mycobacteri
419	39	60.0	1344	2	Q98651	9VIRU	Q98651	rice tungro	492	38.5	59.2	167	2	Q5VHN2_9MYCO	Q5vhn2	mycobacteri
420	39	60.0	1407	2	Q4QIC8	LEIMA	Q4qic8	leishmania	493	38.5	59.2	167	2	Q5VHN5_9MYCFO	Q5vhn5	mycobacteri
421	39	60.0	1411	1	Y0329	HUMAN	O15040	homo sapien	494	38.5	59.2	167	2	Q5VHN6_9MYCO	Q5vhn6	mycobacteri
422	39	60.0	1431	2	Q6ZQE1	MOUSE	Q6zeq1	mus musculus	495	38.5	59.2	167	2	Q5VHN7_9MYCO	Q5vhn7	mycobacteri
423	39	60.0	2014	2	Q54QW2	DICDI	Q54qw2	dictyosteli	496	38.5	59.2	167	2	Q5VHN8_9MYCO	Q5vhn8	mycobacteri
424	39	60.0	2172	2	Q8SVU6	ENCCU	Q8svu6	encephalito	497	38.5	59.2	167	2	Q5VHN9_9MYCO	Q5vhn9	mycobacteri
425	39	60.0	2403	2	O15881	9TRYTP	O15881	trypanosoma	498	38.5	59.2	176	2	Q5QGG2_9MYCO	Q5qgg2	mycobacteri
426	39	60.0	2924	2	Q25733	PLAFA	Q25733	plasmodium	499	38.5	59.2	206	1	SODM_MYCFO	Q59519	mycobacteri
427	39	60.0	3471	1	P0LG	RTSVT	Q91pp5	r genome po	500	38.5	59.2	206	1	SODM_MYCLE	P13367	mycobacteri
428	39	60.0	3473	1	P0LG	RTSVA	Q83034	r genome po								
429	38.5	59.2	129	2	O08238	9MYCO	O08238	mycobacteri								
430	38.5	59.2	129	2	P94916	MYCAB	P94916	mycobacteri								
431	38.5	59.2	129	2	P94934	MYCFO	P94934	mycobacteri								
432	38.5	59.2	129	2	P94935	MYCFO	P94935	mycobacteri								
433	38.5	59.2	129	2	P94936	MYCFO	P94936	mycobacteri								
434	38.5	59.2	129	2	P94959	9MYCO	P94959	mycobacteri								
435	38.5	59.2	129	2	P94961	MYCPR	P94961	mycobacteri								
436	38.5	59.2	129	2	P94969	9MYCO	P94969	mycobacteri								
437	38.5	59.2	135	2	Q6QE21	MYCMA	Q6qe21	mycobacteri								
438	38.5	59.2	137	2	Q581Y6	9MYCO	Q581y6	mycobacteri								
439	38.5	59.2	138	1	SODM	MYCCH	P53644	mycobacteri								
440	38.5	59.2	138	1	SODM	MYCMA	P53645	mycobacteri								
441	38.5	59.2	138	2	Q581Y7	9MYCO	Q581y7	mycobacteri								
442	38.5	59.2	139	2	Q581Y5	9MYCO	Q581y5	mycobacteri								
443	38.5	59.2	147	2	Q5VHN3	9MYCO	Q5vhn3	mycobacteri								
444	38.5	59.2	147	2	Q5VHN4	MYCPR	Q5vhn4	mycobacteri								
445	38.5	59.2	147	2	Q5VHP1	9MYCO	Q5vhp1	mycobacteri								
446	38.5	59.2	147	2	Q5VHP3	9MYCO	Q5vhp3	mycobacteri								
447	38.5	59.2	147	2	Q5VHP4	MYCCH	Q5vhp4	mycobacteri								
448	38.5	59.2	147	2	Q5VHP6	9MYCO	Q5vhp6	mycobacteri								
449	38.5	59.2	147	2	Q5VHP7	MYCAB	Q5vhp7	mycobacteri								
450	38.5	59.2	147	2	Q6QE28	9MYCO	Q6qe28	mycobacteri								
451	38.5	59.2	147	2	Q6QEM5	9MYCO	Q6qem5	mycobacteri								
452	38.5	59.2	147	2	Q6QEM7	9MYCO	Q6qem7	mycobacteri								
453	38.5	59.2	147	2	Q6QEM9	9MYCO	Q6qem9	mycobacteri								
454	38.5	59.2	147	2	Q6QEX0	MYCMA	Q6qex0	mycobacteri								
455	38.5	59.2	147	2	Q6QEX1	MYCMA	Q6qex1	mycobacteri								
456	38.5	59.2	147	2	Q6QEX3	9MYCO	Q6qex3	mycobacteri								
457	38.5	59.2	147	2	Q6QEX4	9MYCO	Q6qex4	mycobacteri								
458	38.5	59.2	147	2	Q6QEX6	MYCRH	Q6qex6	mycobacteri								
459	38.5	59.2	147	2	Q6QEX8	9MYCO	Q6qex8	mycobacteri								
460	38.5	59.2	147	2	Q6QEX0	MYCPR	Q6qex0	mycobacteri								
461	38.5	59.2	147	2	Q6QEX4	MYCNE	Q6qex4	mycobacteri								
462	38.5	59.2	147	2	Q6QEX5	9MYCO	Q6qex5	mycobacteri								
463	38.5	59.2	147	2	Q6QEX6	9MYCO	Q6qex6	mycobacteri								
464	38.5	59.2	147	2	Q6QEX7	9MYCO	Q6qex7	mycobacteri								
465	38.5	59.2	147	2	Q6QEX2	9MYCO	Q6qex2	mycobacteri								
466	38.5	59.2	147	2	Q6QEX5	9MYCO	Q6qex5	mycobacteri								
467	38.5	59.2	147	2	Q6QEX6	9MYCO	Q6qex6	mycobacteri								
468	38.5	59.2	147	2	Q6QF00	9MYCO	Q6qf00	mycobacteri								
469	38.5	59.2	147	2	Q6QF01	9MYCO	Q6qf01	mycobacteri								

ALIGNMENTS

RESULT 1

SEPR_HUMAN	STANDARD	PRT	760 AA.
ID	SEPR_HUMAN	000199; Q86229; Q99998; Q9UID4;	
AC	Q12884; Q00199; Q86229; Q99998; Q9UID4;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Seprase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral membrane serine protease) (170-kDa melanoma membrane-bound gelatinase).		
DE	Gelatinase).		
GN	Name=FAP;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE (ISOFORM 1).		
RC	TISSUE=Fibroblast;		
RX	MEDLINE=94261645; PubMed=7911242;		
RA	Scanlan M.J., Raj B.K.M., Calvo B., Garin-Chesa P., Sanz-Moncasi M.P., Healey J.H., Old L.J., Rettig W.J.;		
RT	"Molecular cloning of fibroblast activation protein alpha, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 91:5657-5661(1994).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT THR-354.		
RC	TISSUE=Melanoma;		
RX	MEDLINE=97389251; PubMed=9247085; DOI=10.1016/S0925-4439(97)00032-X;		
RA	Goldstein L.A., Ghersi G., Pineiro-Sanchez M.L., Salamone M., Yeh Y., Flessate D., Chen W.-T.;		
RT	"Molecular cloning of seprase: a serine integral membrane protease from human melanoma.";		
RL	Biochim. Biophys. Acta 1361:11-19(1997).		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE (ISOFORM 1), AND PROTEIN SEQUENCE OF 220-229;		

461-472 AND 511-518.
 RP TISSUE=Melanoma;
 RC MEDLINE=97218181; PubMed=9065413; DOI=10.1074/jbc.272.12.7595;
 RX Pinciro-Sanchez M.L., Goldstein L.A., Dodi J., Howard L., Yeh Y.,
 RA Chen W.-T.;
 RT "Identification of the 170-kDa melanoma membrane-bound gelatinase
 RT (seprase) as a serine integral membrane protease.";
 RL J. Biol. Chem. 272:7595-7601(1997).
 [4]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
 RC TISSUE=Melanoma;
 RX MEDLINE=20112818; PubMed=10644713; DOI=10.1074/jbc.275.4.2554;
 RA Goldstein L.A., Chen W.-T.;
 RT "Identification of an alternatively spliced seprase mRNA that encodes
 RT a novel intracellular isoform.";
 RL J. Biol. Chem. 275:2554-2559(2000).
 [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
 RC THR-354.
 RX TISSUE=Placenta;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [6]
 RP PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.
 RX MEDLINE=94327249; PubMed=7519584;
 RA Rettig W.J., Su S.B., Fortunato S.R., Scanlan M.J., Raj B.K.M.,
 RA Garin-Chesa P., Healey J.H., Old L.J.;
 RT "Fibroblast activation protein: purification, epitope mapping and
 RT induction by growth factors.";
 RL Int. J. Cancer 58:385-392(1994).
 CC -1- FUNCTION: May have a role in tissue remodeling during development
 CC and wound healing, and may contribute to invasiveness in malignant
 CC cancers.
 CC -1- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
 CC type IV collagen, but not native type I or type IV collagen. Does
 CC not cleave laminin, fibronectin, fibrin or casein.
 CC -1- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
 CC inactive.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell
 CC surface lamellipodia, invadopodia and on shed vesicles.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=L;
 CC IsoId=Q12884-1; Sequence=Displayed;
 CC Note=Major isoform;
 CC Name=2; Synonyms=S, Truncated;
 CC IsoId=Q12884-2; Sequence=VSP_005367;
 CC -1- TISSUE SPECIFICITY: Fibroblast specific.
 CC -1- INDUCTION: In fibroblasts at times and sites of tissue remodeling
 CC during development, tissue repair, and carcinogenesis.
 CC -1- PTM: N-glycosylated.
 CC -1- PTM: The N-terminus may be blocked.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

 CC EMBL; U09278; AB49652.1; -; mRNA.
 CC EMBL; U76833; AAC51668.1; -; mRNA.
 CC EMBL; AF007822; AAF21600.1; -; mRNA.
 CC EMBL; BC026250; AAH26250.1; -; mRNA.
 CC PDB; 1268; X-ray; A=39-757, B=-.
 CC MEROPS; S09.007; -.
 CC Ensembl; ENSG00000078098; Homo sapiens.
 CC HGNC; HGNC:3590; FAP.
 CC MIM; 600403; -.
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0030027; C:lamellipodium; IDA.
 CC GO; GO:0005886; C:plasma membrane; NAS.
 CC GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.
 CC GO; GO:0004222; F:metalloendopeptidase activity; TAS.
 CC GO; GO:0042803; F:protein homodimerization activity; NAS.
 CC InterPro; IPR002471; Pept_S9_AS.
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser_estr.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 CC 3D-structure; Alternative splicing; Direct protein sequencing;
 KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
 KW Signal-anchor; Transmembrane.
 FT TOPO_DOM 1 4
 FT TRANSMEM 5 25
 FT
 FT TOPO_DOM 26 760
 FT ACT_SITE 624 624
 FT ACT_SITE 702 702
 FT ACT_SITE 734 734
 FT CARBOHYD 49 49
 FT CARBOHYD 92 92
 FT CARBOHYD 99 99
 FT CARBOHYD 314 314
 FT CARBOHYD 679 679
 FT VARSPPLIC 1 521
 FT
 FT VARIANT 354 354
 FT CONFLICT 207 207
 FT CONFLICT 229 229
 FT SEQUENCE 760 AA; 87821 MW; A0D34B4801BE07EA CRC64;
 SQ
 Query Match 100.0%; Score 65; DB 1; Length 760;
 Best Local Similarity 100.0%; Pred.No. 0.081; Mismatches 0; Gaps 0;
 Matches 10; Conservative 0; Indels 0;
 QY 1 YALWWSPPNGK 10
 Db 210 YALWWSPPNGK 219
 |||||
 RESULT 2
 Q53TP5 HUMAN
 ID Q53TP5_HUMAN PRELIMINARY; PRT; 760 AA.
 AC Q53TP5;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein FAP.
 GN Name=FAP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Cotton M., Maupin R., Hawkins M., Harkins R.;
 RT "The sequence of Homo sapiens BAC clone RP11-576116";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007750; AAY24205.1; -; Genomic DNA.
 DR GO; GO:0015020; C:membrane; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9 AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estr.
 DR Pfam; PF00930; DPPIV N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 760 AA; 87713 MW; 7FF817B5A4F75142 CRC64;

Query Match 100.0%; Score 65; DB 2; Length 760;
 Best Local Similarity 100.0%; Pred. No. 0.081;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10
 |||||
 DB 210 YALWWSPPNGK 219

RESULT 3
 Q6P7D6 RAT PRELIMINARY; PRT; 350 AA.
 AC Q6P7D6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Fap protein.
 GN Name=Fap;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg K.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Ioshizuki S., Carrincci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Prostate;
 RG NIH MGC Project;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC061713; AAH61713.1; -; mRNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR Pfam; PF00930; DPPIV N; 1.
 DR SEQUENCE 350 AA; 40869 MW; 49FF2B46A9CF0F7F CRC64;
 Query Match 92.3%; Score 60; DB 2; Length 350;
 Best Local Similarity 90.0%; Pred. No. 0.23;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10
 |||||
 DB 210 YALWWSPPNGK 219

RESULT 4
 Q91651 XENLA PRELIMINARY; PRT; 755 AA.
 AC Q91651;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Fibroblast activation factor alpha.
 GN Name=Gene 13;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=96312514; PubMed=8700860; DOI=10.1073/pnas.93.5.1924;
 RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
 RA Remo B.F., Pinder A.;
 RT "The thyroid hormone-induced tail resorption program during Xenopus
 RT laevis metamorphosis";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929 (1996).
 DR EMBL; U41856; AAC59872.1; -; mRNA.
 DR HSSP; P27487; 1PFQ.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9 AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estr.
 DR Pfam; PF00930; DPPIV N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolase; Transmembrane.
 SQ SEQUENCE 755 AA; 86319 MW; E07F41FEB6CB3DF9 CRC64;

Query Match 92.3%; Score 60; DB 2; Length 755;
 Best Local Similarity 90.0%; Pred. No. 0.51;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC embryos from day 7-19 and in new-born mice (P1).
CC -!- SIMILARITY: Belongs to the peptidase S9B family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Y10007; CAA71116.1; -; mRNA.
CC EMBL; BC019190; AAH19190.1; -; mRNA.
CC HSSP; P27487; 1PFQ.
CC MEROPS; S09.007;
CC ENSEMBL; ENSMUSG00000000392; Mus musculus.
CC MGI; MGI:109608; Fap.
CC GO; GO:0016021; C: integral to membrane; TAS.
CC InterPro; IPR002471; Pept_S9_AS.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_eatrs.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase_S9; 1.
CC PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
CC Alternative splicing; Glycoprotein; Hydrolase; Protease;
CC Serine protease; Signal-anchor; Transmembrane.
CC TOPO_DOM 1 4 Cytoplasmic (Potential).
CC TRANSMEM 5 25 Signal-anchor for type II membrane
CC protein (Potential).
CC TOPO_DOM 26 761 Extracellular (Potential).
CC ACT_SITE 624 624 Charge relay system (By similarity).
CC ACT_SITE 702 702 Charge relay system (By similarity).
CC ACT_SITE 734 734 Charge relay system (By similarity).
CC CARBOHYD 49 49 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 92 92 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 227 227 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
CC VARSPLIC 31 63 Missing (in isoform 3).
CC VARSPLIC 31 35 Missing (in isoform 2).
CC CONFLICT 737 737 S -> L (in Ref. 2).
CC SEQUENCE 761 AA; 87945 MW; 9174C3AEDA213B25 CRC64;
SQ
Query Match 92.3%; Score 60; DB 1; Length 761;
Best Local Similarity 90.0%; Pred. No. 0.51;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YALWVSPNGK 10
DB 210 YALWVSPDCK 219

CC RESULT 6
CC Q8R492 RAT PRELIMINARY; PRT; 761 AA.
CC ID Q8R492 RAT PRELIMINARY; PRT; 761 AA.
CC AC Q8R492;
CC DT 01-JUN-2002 (TrEMBLrel. 21, Created)
CC DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
CC DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
CC DE Fibroblast activation protein alpha subunit.
CC GN NamesFap;
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
CC OC Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP NUCLEOTIDE SEQUENCE.
CC RA Zuckeswerdt K., Park J.E.;
CC RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC EMBL; AF493782; AA011677.1; -; mRNA.

CC 1 YALWVSPNGK 10
CC 207 YATWVSPNGK 216

CC RESULT 5
CC SEPR_MOUSE STANDARD; PRT; 761 AA.
CC AC P97321;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-MAY-2005 (Rel. 47, Last annotation update)
CC DE Sepsinase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
CC membrane serine protease).
CC GN Name=Fap;
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
CC OC Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).
CC RC STRAIN=BALB/c; TISSUE=Embryo;
CC RX MEDLINE=97284459; PubMed=9139873;
CC RA Niedermeyer J., Scanlan M.J., Garin-Chesa P., Daiber C., Fiebig H.H.,
CC Old L.J., Rettig W.J., Schnapp A.;
CC RT "Mouse fibroblast activation protein: molecular cloning, alternative
CC splicing and expression in the reactive stroma of epithelial
CC cancers.";
CC RT Int. J. Cancer 71:383-389(1997).
CC RN [2]
CC RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
CC RC TISSUE=Mammary gland;
CC RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
CC Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
CC Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
CC Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
CC Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
CC Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
CC Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
CC Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smailus D.E.,
CC Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
CC RT "Generation and initial analysis of more than 15,000 full-length human
CC and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC RL FUNCTION: May have a role in tissue remodeling during development
CC and wound healing, and contribute to invasiveness in malignant
CC cancers.
CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
CC type IV collagen, but not native type I or type IV collagen. Does
CC not cleave laminin, fibronectin, fibrin or casein.
CC -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
CC inactive (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P97321-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P97321-2; Sequence=VSP_005368;
CC Name=3;
CC IsoId=P97321-3; Sequence=VSP_005369;
CC -!- TISSUE SPECIFICITY: Detected in fibroblasts, in placenta, uterus,

DR HSP; P27487; 1PFQ.
DR MEROPS; S09.007; -;
DR Ensembl; ENSRNOG00000005679; Rattus norvegicus.
DR RGD; 621253; Fap.

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:hydrolyase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 761 AA; 87843 MW; EB214BBE4BB14F2C CRC64;

Query Match 92.3%; Score 60; DB 2; Length 761;
Best Local Similarity 90.0%; Pred. No. 0.51;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWWSNGK 10
|||||:|
DB 210 YALWWSPDGK 219

RESULT 7

ID Q75S82_AGKHB PRELIMINARY; PRT; 751 AA.
AC Q75S82;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Dipeptidylpeptidase 4b (EC 3.4.14.5).
GN Name=DPP4b;

OS Agkistrodon halya brevicaudus (Korean slamosa snake) (Gloydus halya brevicaudus).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=259325;
RN [1]

RP TISSUE=Venom gland.

RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158225; BAD06333.1; -; mRNA.

DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase S9; 1.

DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.

SQ SEQUENCE 751 AA; 86155 MW; DAF219B4FEE7629A CRC64;

Query Match 84.6%; Score 55; DB 2; Length 751;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSNGK 10
|||||:|
DB 207 ALWWSNGR 215

RESULT 8

Q75S83_AGKHB
ID Q75S83_AGKHB PRELIMINARY; PRT; 751 AA.

AC Q75S83;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Dipeptidylpeptidase 4a (EC 3.4.14.5).
GN Name=DPP4a;
OS Agkistrodon halya brevicaudus (Korean slamosa snake) (Gloydus halya brevicaudus).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;

OC Viperidae; Crotalinae; Gloydus.

OX NCBI_TaxID=259325;
RN [1]

RP TISSUE=Venom gland.

RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB158224; BAD06332.1; -; mRNA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0004177; F:aminopeptidase activity; IEA.

DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.

DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR InterPro; IPR002471; Pept S9 AS.

DR InterPro; IPR001375; Peptidase S9.

DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser_estrs.

DR Pfam; PF00930; DPPIV_N; 1.

DR Pfam; PF00326; Peptidase S9; 1.

DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.

KW Hydrolase; Transmembrane.

SQ SEQUENCE 751 AA; 86127 MW; 6EA3F2D829B597D4 CRC64;

Query Match 84.6%; Score 55; DB 2; Length 751;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSNGK 10
|||||:|
DB 207 ALWWSNGR 215

RESULT 9

Q68K44_9HEPC
ID Q68K44_9HEPC PRELIMINARY; PRT; 659 AA.

AC Q68K44;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,

RA Kleiner D., Holman S., Augenbraun M., Taylor J.;

RT "Sequence Analysis of Hepatitis C Virus Replication Functions in HCV/HIV Coinfected Subjects."

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY685622; AA194269.1; -; Genomic_RNA.

DR SMR; Q68K44; 1-552.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO:0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR004109; Peptidase_S29.
 DR Pfam: PF02907; HCV NS3; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
 KW Structural protein.
 FT NON_TER 1
 FT NON_TER 659
 SQ SEQUENCE 659 AA; 70398 MW; 2ACAEHLF20ADF453 CRC64;

Query Match 83.1%; Score 54; DB 2; Length 659;
 Best Local Similarity 88.9%; Pred. No. 4;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSFNGK 10
 DB 642 ALWWSFNGE 650

RESULT 10
 Q52181 CHICK PRELIMINARY; PRT; 759 AA.
 AC Q52181;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=RCJMB04.29921;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=CB; TISSUE=Bursa;
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
 RA Fiedler P., Kutter S., Biagodevski A., Kostovska D., Koter M.,
 RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
 RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
 RT gene function analysis."
 RL Genome Biol. 6:R6-R6(2005).
 DR EMBL; AJ720903; CAG32562.1; -; mRNA.
 DR GO:0016020; C:membrane; IEA.
 DR GO:0004177; F:aminopeptidase activity; IEA.
 DR GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO:0016787; F:hydrolase activity; IEA.
 DR GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002471; Pept S9 AS.
 DR InterPro: IPR001375; Peptidase_S9.
 DR InterPro: IPR002469; Peptidase_S9B.
 DR InterPro: IPR000379; Ser. esters.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 759 AA; 86343 MW; B9A0C38CCB644AC CRC64;

Query Match 81.5%; Score 53; DB 2; Length 759;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSFNG 9

DB 211 ALWWSFNG 218

RESULT 11
 Q5R7G7 PONPY PRELIMINARY; PRT; 765 AA.
 ID Q5R7G7;
 AC Q5R7G7;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein DKFZp469P1419.
 GN Name=DKFZp469P1419;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
 OC Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Kidney;
 RG The German cDNA Consortium;
 RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; CR860150; CAH92293.1; -; mRNA.
 DR SMR; Q5R7G7; 38-765.
 DR GO:0016020; C:membrane; IEA.
 DR GO:0004177; F:aminopeptidase activity; IEA.
 DR GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO:0016787; F:hydrolase activity; IEA.
 DR GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002471; Pept S9 AS.
 DR InterPro: IPR001375; Peptidase_S9.
 DR InterPro: IPR002469; Peptidase_S9B.
 DR InterPro: IPR000379; Ser. esters.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 765 AA; 88038 MW; 6F81ECD98C3D2397 CRC64;

Query Match 81.5%; Score 53; DB 2; Length 765;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSFNG 9
 DB 212 ALWWSFNG 219

RESULT 12
 DPP4_HUMAN
 ID DPP4_HUMAN STANDARD; PRT; 766 AA.
 AC P27487;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
 DE IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase
 DE complexing protein 2) (ADAP) (Contains: Dipeptidyl peptidase 4
 DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
 DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)).
 GN Name=DPP4; Synonyms=ADCF2, CD26;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Liver;

RA MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT "Molecular cloning and sequence analysis of human dipeptidyl peptidase
RT IV, a serine proteinase on the cell surface.";
RL Biochim. Biophys. Acta 1131:333-336(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=95012454; PubMed=7927537;
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT "Genomic organization, exact localization, and tissue expression of
RT the human CD26 (dipeptidyl peptidase IV) gene.";
RL Immunogenetics 40:331-338(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral blood;
RX MEDLINE=92325476; PubMed=1352530;
RA Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT "Cloning and functional expression of the T cell activation antigen
RT CD26.";
RL J. Immunol. 149:481-486(1992).
RN [4]
RP ERRATUM.
RX MEDLINE=93171637; PubMed=8094732;
RA Tanaka T.;
RL J. Immunol. 150:2090-2090(1993).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marta M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 1-551.
RC TISSUE=Colon;
RX MEDLINE=92165847; PubMed=1347043;
RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,
RA Barbat A.;
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT levels during cell differentiation.";
RL J. Biol. Chem. 267:4824-4833(1992).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 545-766.
RC TISSUE=Colon;
RX MEDLINE=91024044; PubMed=1977364;
RA Darmoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidylpeptidase IV and assignment of the gene locus DPP4 to
RT chromosome 2.";
RL Ann. Hum. Genet. 54:191-197(1990).
RN [8]
RP NUCLEOTIDE SEQUENCE OF 1-31.

RX MEDLINE=96067599; PubMed=7487939;
RA Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT "Human dipeptidyl peptidase IV gene promoter: tissue-specific
RT regulation from a TATA-less GC-rich sequence characteristic of a
RL housekeeping gene promoter.";
RL Biochem. J. 311:835-843(1995).
RN [9]
RP PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
RX PubMed=1677636;
RA Garovel J.P., Ferrero A., Chambraud L., Rigal A., Bonicel J.,
RA Maroux S.;
RT "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human
RT small intestine and colon.";
RL Gastroenterology 101:618-625(1991).
RN [10]
RP PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=93210469; PubMed=8096237; DOI=10.1084/jem.177.4.1135;
RA Morrison M.E., Vijayaaradhi S., Engelstein D., Albino A.P.,
RA Houghton A.N.;
RT "A marker for neoplastic progression of human melanocytes is a cell
RT surface ectopeptidase.";
RL J. Exp. Med. 177:1135-1143(1993).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH
RP INHIBITOR AND HOMODIMERIZATION.
RX PubMed=12832764; DOI=10.1107/S0907444903010059;
RA Oefner C., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R.,
RA Dale G.E.;
RT "High-resolution structure of human apo dipeptidyl peptidase IV/CD26
RT acetyl-L-2-cyano-(S)-pyrrolidine.";
RL Acta Crystallogr. D 59:1206-1212(2003).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=1246248; DOI=10.1016/S0006-291X(03)00258-4;
RA Hiramatsu H., Kyono K., Higashiyama Y., Fukushima C., Shima H.,
RA Sugiyama S., Inaka K., Yamamoto A., Shimizu R.;
RT "The structure and function of human dipeptidyl peptidase IV,
RT possessing a unique eight-bladed beta-propeller fold.";
RL Biochem. Biophys. Res. Commun. 302:849-854(2003).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12483204; DOI=10.1038/nsb882;
RA Rasmussen H.B., Branner S., Wiberg F.C., Wagtmann N.;
RT "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex
RT with a substrate analog.";
RL Nat. Struct. Biol. 10:19-25(2003).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;
RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;
RT "Structural basis of proline-specific exopeptidase activity as
RT observed in human dipeptidyl peptidase-IV.";
RL Structure 11:947-959(2003).
RN [15]
RP FUNCTION: Removes N-terminal dipeptides sequentially from
RP polypeptides having unsubstituted N-termini provided that the
RP penultimate residue is proline. Plays a role in T cell activation.
CC CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC SUBUNIT: Homodimer or heterodimer with seprase (FAP).
CC SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt
CC cells of the small intestine as well as in the mature villous
CC cells. Expressed at very low levels in the colon.
CC PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; U13735; AAB06046.1; -, Genomic DNA.
DR EMBL; U13710; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13711; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13712; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13713; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13714; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13715; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13716; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13717; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13718; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13719; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13720; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13721; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13722; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13723; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13724; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13725; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13726; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13727; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13728; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13729; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13730; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13731; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13732; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13733; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; U13734; AAB06046.1; JOINED; Genomic DNA.
DR EMBL; M74777; AAB06046.1; -, mRNA.
DR EMBL; BC013329; AAB13329.2; -, mRNA.

Query Match 81.5%; Score 53; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWMSPPNG 9
|||||||
DB 213 ALWMSPPNG 220

RESULT 13
DPP4_PIG STANDARD; PRT; 766 AA.
ID P2241; Q866G2;
AC P2241; Q866G2;
DT 01-AUG-1991 (Rel. 19, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (Contains: Dipeptidyl peptidase 4
DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)).
GN Names: DPP4; Synonyms: CD26;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus
OX NCBI_TaxID=9823;
[1]
RN NUCLEOTIDE SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RC PubMed=14719797; DOI=10.1515/BC.2003.172;
RA Baer J., Weber A., Hoffmann T., Stork J., Wermann M., Wagner L.,
RA Aust S., Gerhartz B., Demuth H.-U.;
RA "Characterisation of human dipeptidyl peptidase IV expressed in Pichia
RT pastoris. A structural and mechanistic comparison between the
RT recombinant human and the purified porcine enzyme.";
RL Biol. Chem. 384:1553-1563(2003).
RN [2]

NUCLEOTIDE SEQUENCE OF 2-67.

RP TISSUE=Kidney;
RC MEDLINE=94093209; PubMed=7903569;
RA Thomsen P.D., Qvist H., Marklund L., Andersson L., Sjoestrom H.,
RA Noren O.;
RT "Assignment of the dipeptidylpeptidase IV (DPP4) gene to pig
RT chromosome 15q21.";
RL Mamm. Genome 4:604-607(1993).
[3]
RN PROTEIN SEQUENCE OF 38-71.
RP TISSUE=Kidney;
RC MEDLINE=91273813; PubMed=1675855;
RX Seidl R., Mann K., Schaeffer W.;
RA "N-terminal amino-acid sequence of pig kidney dipeptidyl peptidase IV
RT solubilized by autolysis.";
RL Biol. Chem. Hoppe-Seyler 372:213-214(1991).
[4]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 39-766, N-GLYCOSYLATION
RP SITES, AND HOMODIMERIZATION.
RX PubMed=12690074; DOI=10.1073/pnas.0230620100;
RA Engel M., Hoffmann T., Wagner L., Wermann M., Heiser U.,
RA Kieferauer R., Huber R., Bode W., Demuth H.-U., Brandstetter H.;
RT "The crystal structure of dipeptidyl peptidase IV (CD26) reveals its
RT functional regulation and enzymatic mechanism.";
CC Proc. Natl. Acad. Sci. U.S.A. 100:5063-5068(2003).
CC -|- FUNCTION: Removes N-terminal dipeptides sequentially from the
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline.
CC -|- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -|- SUBUNIT: Homodimer.
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -|- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -|- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL; AY198323; AAO43404.1; -, mRNA.
DR EMBL; X73276; CAAS1717.1; -, mRNA.
DR EIR; I47134; I47134.
DR PDB; 1ORV; X-ray; A/B/C/D=39-766.
DR PDB; 1ORW; X-ray; A/B/C/D=39-766.
DR SMR; P2241; 39-766.
DR MEROPS; S09.003; -.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
DR 3D-structure; Amino-peptidase; Direct protein sequencing; Glycoprotein;
KW Dipeptidyl peptidase 4 membrane form.
KW Hydroxylase; Protease; Serine protease.
FT CHAIN 1 766
FT CHAIN 38 766
FT CHAIN 1 6
FT TOPO_DOM 7 27
FT TRANSMEM
FT TOPO_DOM 28 766
FT ACT_SITE 630 630
FT ACT_SITE 708 708
FT ACT_SITE 740 740
FT CARBOHYD 85 85
FT CARBOHYD 92 92
FT CARBOHYD 150 150
FT CARBOHYD 179 179
FT N-linked (GlcNAc...) (By similarity).
FT N-linked (GlcNAc...) (Potential).

FT CARBOHYD 219 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 229 N-linked (GlcNAc...)
 FT CARBOHYD 279 N-linked (GlcNAc...)
 FT CARBOHYD 321 N-linked (GlcNAc...)
 FT CARBOHYD 393 N-linked (GlcNAc...)
 FT CARBOHYD 685 N-linked (GlcNAc...)
 FT DISULFID 395
 FT DISULFID 444
 FT DISULFID 454
 FT DISULFID 649
 FT CONFLICT 32
 SQ SEQUENCE 766 AA; 88242 MW; 8800D520BAEA856D CRC64;

Query Match 81.5%; Score 53; DB 1; Length 766;

Best Local Similarity 100.0%; Pred. No. 6.7; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0;

Oy 2 ALWWSPNG 9
 |||||
 Db 213 ALWWSPNG 220

RESULT 14
 Q53TN1 HUMAN
 ID Q53TN1_HUMAN PRELIMINARY; PRT; 766 AA.
 AC Q53TN1_1
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Hypothetical protein DPP4.
 GN Name=DPP4;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 ON NCBI_TaxID=9606;
 RP NUCLEOTIDE SEQUENCE.
 RA Ozanich A., Stoneking T., Hawkins M., Sapetti L.;
 RT "The sequence of Homo sapiens BAC clone RP11-178A14";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBSJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBSJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBSJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AC008063; AAX93179.1; -; Genomic_DNA.
 DR SMR; Q53TN1; 39-766.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estr.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 766 AA; 88279 MW; 5FB4A2C662D6117 CRC64;

Query Match 81.5%; Score 53; DB 2; Length 766;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 ALWWSPNG 9
 |||||
 Db 213 ALWWSPNG 220

RESULT 15
 DPP4_RAT
 ID _DPP4_RAT STANDARD; PRT; 767 AA.
 AC P14740;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-2005 (Rel. 46, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV) (T-cell activation antigen CD26) (gp110 glycoprotein) (Bile canalculus domain-specific membrane glycoprotein) [Contains: Dipeptidyl peptidase 4 membrane form (dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl peptidase IV 60 kDa soluble form)].
 GN Name=Dpp4; Synonyms=Cd26;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
 RX MEDLINE=89123496; PubMed=2563382;
 RA Ogata S., Misumi Y., Ikehara Y.;
 RT "Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA and identification of the NH2-terminal signal sequence as the membrane-anchoring domain";
 RL J. Biol. Chem. 264:3596-3601 (1989).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=88068516; PubMed=3479775;
 RA Hong W., Doyle D.;
 RT "cDNA cloning for a bile canalculus domain-specific membrane glycoprotein of rat hepatocytes";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966 (1987).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 1-40.
 RX MEDLINE=89034185; PubMed=3182821;
 RA Hong W.J., Doyle D.;
 RT "Membrane orientation of rat gp110 as studied by in vitro translation";
 RL J. Biol. Chem. 263:16892-16898 (1988).
 RN [4]
 RP PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY.
 RX MEDLINE=90228896; PubMed=1970322;
 RA McCaughan G.W., Wickson J.E., Creswick P.F., Gorrell M.D.;
 RT "Identification of the bile canalicular cell surface molecule GP110 as the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue distribution, purification and N-terminal amino acid sequence";
 RL Hepatology 11:534-544 (1990).
 RN [5]
 RP PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630; SER-631; TRY-632 AND GLY-633.
 RX TISSUE=Kidney;
 RX MEDLINE=94128239; PubMed=7905271;
 RA Iwaki-Egawa S., Watanabe Y., Fujimoto Y.;
 RT "N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl peptidase IV";
 RL Biol. Chem. Hoppe-Seyler 374:973-975 (1993).
 RN [6]
 RP PROTEIN SEQUENCE OF 624-648.
 RX MEDLINE=92190188; PubMed=1347701;
 RA Ogata S., Misumi Y., Teuji E., Takami N., Oda K., Ikehara Y.;
 RT "Identification of the active site residues in dipeptidyl peptidase IV by affinity labeling and site-directed mutagenesis";
 RL Biochemistry 31:2582-2587 (1992).
 RN [7]

DR Pfam: PF00930; DPPIV N; 1.
 DR Pfam: PF00326; Peptidase S9; 1.
 DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.
 KW Aminopeptidase; Complete proteome; Hydrolase; Transmembrane.
 SQ SEQUENCE 817 AA; 93991 MW; 9789F01F4F967FD CRC64;

Query Match 81.5%; Score 53; DB 2; Length 817;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPNG 9
 |||||
 Db 247 ALWWSPNG 254

RESULT 17

ID Q5SH13 CRYNE PRELIMINARY; PRT; 883 AA.
 AC Q5SH13
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=CNBN1370;
 OS Cryptococcus neoformans var. neoformans B-3501A.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=283643;
 (1)

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=B-3501A;
 RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
 Wickes B.L., Fu J., Davis R.W.;
 RA "Cryptococcus neoformans serotype D sequencing."
 RT Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.

DR EMBL; AA01000066; EAL17310.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estrs.
 DR Pfam; PF00930; DPPIV N; 1.
 DR PROSITE; PS00708; PRO-ENDOPEP_SER; UNKNOWN 1.
 KW Hydrolase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 883 AA; 99072 MW; 2B66FFA3F54445F1 CRC64;

Query Match 81.5%; Score 53; DB 2; Length 883;
 Best Local Similarity 88.9%; Pred. No. 7.8;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPNGK 10
 |||||
 Db 303 ALWWSPDGK 311

RESULT 18

ID Q5K721 CRYNE PRELIMINARY; PRT; 883 AA.
 AC Q5K721
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Dipeptidyl-peptidase and tripeptidyl-peptidase, putative.
 GN ORFNames=CNBN1410;
 OS Cryptococcus neoformans var. neoformans JEC21.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;

OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=214694;
 (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=JEC21;
 RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Uterback T.,
 Van Aken S., Fraser C.;
 RA Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 (2)

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=JEC21;
 RC PubMed=15653466; DOI=10.1126/science.1103773;
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
 Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
 Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
 D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
 Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
 Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
 Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
 Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
 Suh B.B., Tenney A., Uterback T.R., Wickes B.L., Wortman J.R.,
 Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
 Fraser C.M., Hyman R.W.;
 RA "The genome of the basidiomycetous yeast and human pathogen
 Cryptococcus neoformans."
 RT Science 307:1321-1324(2005).
 RL EMBL; AE017356; AA047088.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estrs.
 DR Pfam; PF00930; DPPIV N; 1.
 DR PROSITE; PS00708; PRO-ENDOPEP_SER; UNKNOWN 1.
 KW Complete proteome; Hydrolase; Transmembrane.

SQ SEQUENCE 883 AA; 99072 MW; 2B66FFA3F54445F1 CRC64;

Query Match 81.5%; Score 53; DB 2; Length 883;
 Best Local Similarity 88.9%; Pred. No. 7.8;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPNGK 10
 |||||
 Db 303 ALWWSPDGK 311

RESULT 19

ID Q75C44 ASHGO PRELIMINARY; PRT; 872 AA.
 AC Q75C44;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ACR073Cp.
 GN Name=ACR073C;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
 OX NCBI_TaxID=33169;
 (1)

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=ATCC 10895;
 RC PubMed=15001715; DOI=10.1126/science.1095781;
 RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
 Mohr C., Poehmann R., Luedi P., Choi S., Wing R.A., Flavler A.,
 Gaffney T.D., Philippsen P.;
 RA "The Ashbya gossypii genome as a tool for mapping the ancient
 Saccharomyces cerevisiae genome.";

```
Science 304:304-307(2004).
RL EMBL; AR016816; AAS51299.1; -, Genomic_DNA.
DR AGD; ACR073C; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0004177; F-aminopeptidase activity; IEA.
DR GO; GO:0004274; F-dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F-hydrolase activity; IEA.
DR GO; GO:0004287; F-prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Complete proteome; Hydrolase; Transmembrane.
SQ SEQUENCE 872 AA; 98529 MW; 4E7FEA3728EE1E07 CRC64;

Query Match 80.0%; Score 52; DB 2; Length 872;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSFNGK 10
DB 305 ALWWSFNGK 313

RESULT 20
Q6PG25 XENLA
ID Q6PG25 XENLA PRELIMINARY; PRT; 439 AA.
AC Q6PG25;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC64382 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner B., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative."

Dev. Dyn. 225:384-391(2002).
RL NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC057298; AAH57298.1; -, mRNA.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0004274; F-dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
SQ SEQUENCE 439 AA; 50314 MW; 14212078E30D8A18 CRC64;

Query Match 78.5%; Score 51; DB 2; Length 439;
Best Local Similarity 80.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWWSFNGK 10
DB 195 YALWWSFNGK 204

RESULT 21
Q82C69 STRAW
ID Q82C69 STRAW PRELIMINARY; PRT; 404 AA.
AC Q82C69;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SAV5485;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC73197.1; -, Genomic DNA.
SQ SEQUENCE 404 AA; 43046 MW; 76CA609577776938 CRC64;

Query Match 75.4%; Score 49; DB 2; Length 404;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSFNGK 9
DB 349 YALWWSFNGK 357

RESULT 22
Q641D6 XENLA
ID Q641D6 XENLA PRELIMINARY; PRT; 737 AA.
AC Q641D6;
```

25-OCT-2004 (TrEMBLrel. 28, Created)
 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 LOC397888 protein.
 Name=LOC397888;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.",
 RL Dev. Dyn. 225:384-391 (2002).
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC082403; AA82403.1; -; mRNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0016787; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0005058; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR Pfam; PF00930; DPPIV_N; I.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 DR HydroLase; Transmembrane.
 SQ SEQUENCE 737 AA; 84184 MW; 1E08CF94EDFBC4B CRC64;
 Query Match 75.4%; Score 49; DB 2; Length 737;
 Best Local Similarity 87.5%; Pred. No. 28;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YALWWSNP 8
 DB 195 YALWWSNP 202

RESULT 23
 P70092_XENLA PRELIMINARY; PRT; 748 AA.
 AC P70092;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Dipeptidyl-peptidase IV (EC 3.4.14.5).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 NCBI_TaxID=8355;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Skin;
 RX MEDLINE=97390076; PubMed=9249015;
 RA Vlasak R., Vilas U., Strobl B., Kreil G.;
 RA "CDNA cloning and expression of secreted xenopus laevis dipeptidyl
 RT aminopeptidase IV.",
 RL Eur. J. Biochem. 247:107-113 (1997).
 DR EMBL; Y08932; CAA70136.1; -; mRNA.
 DR HSSP; P27487; 1PFQ.
 DR MEROPS; S09.003; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0005058; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_eetra.
 DR Pfam; PF00930; DPPIV_N; I.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 DR HydroLase; Transmembrane.
 SQ SEQUENCE 748 AA; 85587 MW; 31CCF61B25104E5B CRC64;
 Query Match 75.4%; Score 49; DB 2; Length 748;
 Best Local Similarity 87.5%; Pred. No. 28;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YALWWSNP 8
 DB 206 YALWWSNP 213

RESULT 24
 Q6CFX2_YARLI PRELIMINARY; PRT; 868 AA.
 AC Q6CFX2;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to tr|Q96VT7|Aspergillus niger Dipeptidyl aminopeptidase type
 DE IV.
 GN OrderedLocusNames=YAL10B02838g;
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 NCBI_TaxID=4952;
 [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CLIB 122 / E 150;
 RX PubMed=15229592; DOI=10.1038/nature02579;
 RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Falla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
 RA Boissiere A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,

RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicraud A., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-P., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J.,
RA Wincker P., Souciet J.-L.,
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
DR EMBL; CR382128; CAG82658.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Aminopeptidase; Complete proteome; Lipid degradation.
SQ SEQUENCE 868 AA; 97775 MW; A6C4244DFC15F3F2 CRC64;

Query Match 75.4%; Score 49; DB 2; Length 868;
Best Local Similarity 77.8%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
DB 314 ALWSPDQG 322
|||||

RESULT 25
UBIG RHIL0 STANDARD; PRT; 249 AA.
AC Q98G87;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
GN Name:UBIG, OrderedLocusNames=mir3442;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
CC Ubig/COO3 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; BA000012; BAB50329.1; -; Genomic_DNA.
DR HAMAP; MF_00472; -; 1.
DR InterPro; IPR003333; CMAS.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bd.
DR InterPro; IPR010233; Ubig_mtfrase.
DR Pfam; PF02353; CMAS; 1.
DR TIGRFAMs; TIGR01983; Ubig; 1.
KW Complete proteome; Methyltransferase; Transferase;
KW Ubiquinone biosynthesis.
SQ SEQUENCE 249 AA; 27318 MW; F8216C9041F0AC5A CRC64;

Query Match 73.8%; Score 48; DB 1; Length 249;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
DB 22 AEWNPNGK 30
|||||

RESULT 26
Q4S309 TETNG PRELIMINARY; PRT; 704 AA.
ID Q4S309;
AC Q4S309;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAP14756, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00024844001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Maudeli E., Bouteau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014756; CAG04973.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 704 AA; 79459 MW; 61F46625DF618619 CRC64;

Query Match 73.8%; Score 48; DB 2; Length 704;
Best Local Similarity 87.5%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWWSPNGK 10
DB 105 LWWSPGK 112
|||||

```

RESULT 27
Q60J70 CAEBR PRELIMINARY; PRT; 794 AA.
AC Q60J70;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG24661.
GN Name=CBG24661;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
[1]
RN NUCLEOTIDE SEQUENCE.
RP The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC0100344; CAES6836.1; -; Genomic_DNA.
FT NON TER 1
SQ SEQUENCE 553 AA; 63057 MW; 0F7BA4D3C73E2048 CRC64;

Query Match 72.3%; Score 47; DB 2; Length 553;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALMWSPNG 9
DB 410 AIWHPNG 417

RESULT 29
DPP4_MOUSE
ID DPP4_MOUSE STANDARD; PRT; 760 AA.
AC P28843;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
IV) (T-cell activation antigen CD26) (Thymocyte-activating molecule)
DE (TMAM) [Contains: dipeptidyl peptidase 4 membrane form (Dipeptidyl
peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form]
DE (Dipeptidyl peptidase IV soluble form)].
GN Name=Dpp4; Synonyms=Cd26;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Swiss; TISSUE=Thymus;
RX MEDLINE=92129286; PubMed=1370813;
RA Marguet D.A.; Bernard A.-M., Vivier I., Darmoul D., Naquet P.,
Pierres M.;
RA "cDNA cloning for mouse thymocyte-activating molecule. A
multifunctional ecto-dipeptidyl peptidase IV (CD26) included in a
subgroup of serine proteases";
RT J. Biol. Chem. 267:2200-2208(1992).
[2]
RN SEQUENCE REVISION.
RA Marguet D.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RX STRAIN=B10.A; TISSUE=Liver;
RX MEDLINE=95092780; PubMed=7999781;
RA Bernard A.-M., Mattei M.-G., Pierres M., Marguet D.;
RT "Structure of the mouse dipeptidyl peptidase IV (CD26) gene.";
RL Biochemistry 33:15204-15214(1994).
[4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
Nikaido I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

```

```

Query Match      72.3%; Score 47; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. NO. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSNP 8
          |||||
Db      207 ALWWSNP 213

```


[1] NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral blood;
RX MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;
RA Nishimura Y., Miyazawa T., Ikeda Y., Izumiya Y., Nakamura K., Sato E.,
RA Mikami T., Takahashi E.;
RT "Molecular cloning and sequencing of a cDNA encoding the feline T-cell
RL Immunogenetics 50:366-368(1999).
CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline (By similarity).
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form (By similarity).
CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing (By similarity).
CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; A8023952; BAA92344.1; -; mRNA.
DR HSP; P27487; 1PQ.
DR SMR; Q9N217; 38-765.
DR MEROPS; S09.003; -.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00336; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SRR; 1.
DR AminoPeptidase; Glycoprotein_Hydrolase; Protease; Serine protease;
KW Signal-anchor; Transmembrane.
KW CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.
FT FT 38 765 Dipeptidyl peptidase 4 soluble form (By
FT FT 1 765 similarity).
FT FT 6 Cytoplasmic (Potential).
FT FT 29 Signal-anchor for type II membrane
FT FT protein (Potential).
FT FT 765 Extracellular (Potential).
FT FT 629 Charge relay system (By similarity).
FT FT 707 Charge relay system (By similarity).
FT FT 739 Charge relay system (By similarity).
FT FT 84 N-linked (GlcNAc...) (By similarity).
FT FT 91 N-linked (GlcNAc...) (By similarity).
FT FT 149 N-linked (GlcNAc...) (By similarity).
FT FT 178 N-linked (GlcNAc...) (Potential).
FT FT 228 N-linked (GlcNAc...) (By similarity).
FT FT 280 N-linked (GlcNAc...) (By similarity).
FT FT 320 N-linked (GlcNAc...) (By similarity).
FT FT 330 N-linked (GlcNAc...) (Potential).
FT FT 331 N-linked (GlcNAc...) (Potential).
FT FT 519 N-linked (GlcNAc...) (By similarity).
FT FT 684 N-linked (GlcNAc...) (By similarity).
FT FT 384 By similarity.
FT FT 443 By similarity.
FT FT 453 By similarity.
FT FT 648 By similarity.
SQ SEQUENCE 765 AA; 38FC98A22B175D9 CRC64;
Query Match 72.3%; Score 47; DB 1; Length 765;
Best Local Similarity 87.5%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ALWWSPPNG 9

Db 212 ALWWSPPKG 219
RESULT 32
QAHVM6 GIBZE PRELIMINARY; PRT; 779 AA.
ID QAHVM6;
AC QAHVM6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFName=FG10382.1;
OS Gibberella zeae PH-1.
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1] NUCLEOTIDE SEQUENCE.
RP STRAIN=PH-1;
RC Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kella C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tsefaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence."
RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data
CC EMBL; AAC0100457; EAA75308.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 779 AA; 87757 MW; 3EC9D468C633F192 CRC64;
Query Match 72.3%; Score 47; DB 2; Length 779;
Best Local Similarity 77.8%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 ALWWSPPNGK 10
Db 208 ALWWSPPDSK 216
RESULT 33
Q6CNM1 KLJLA PRELIMINARY; PRT; 829 AA.
ID Q6CNM1;
AC Q6CNM1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome E of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
GN OrderedLocNames=KLJA0E114629;
OS Kluyveromyces lactis (yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

```
OX NCBI_TaxID=28985;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CBS 2359 / IFO 1267 / NRRL Y-1140 / WM37;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Coffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin J.-M., Beckerich J.-W., Beyne E., Bleykasten C.,
RA Boigranme A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dunazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozul R., Lemaire M., Leaur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Fellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR382125; CAG99555.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Complete proteome; Hydrolase; Transmembrane.
SQ SEQUENCE 829 AA; 94226 MW; 733BB79B040C8D9 CRC64;

Query Match 72.3%; Score 47; DB 2; Length 829;
Best Local Similarity 87.5%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPTG 9
DB 260 ALWWSPTG 267

RESULT 34
Q18253 CAEBL
ID Q18253 CAEBL PRELIMINARY; PRT; 829 AA.
AC Q18253;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein dpf-2;
GN Name=dpf-2; ORFNames=C27C12.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN NUCLEOTIDE SEQUENCE.
RP Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RA Thomas K.;
RA STRAIN=Bristol N2;
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
```

```
DR EMBL; Z69883; CAA93743.1; -; Genomic_DNA.
DR PIR; T19514; T19514.
DR HSSP; P27487; 1PFO.
DR Ensembl; C27C12.7; Caenorhabditis elegans.
DR WormBase; WBGene00001055; dpf-2.
DR WormPep; C27C12.7; CE05324.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 829 AA; 94387 MW; B3F76F6DC12E44A5 CRC64;

Query Match 72.3%; Score 47; DB 2; Length 829;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPTGK 10
DB 265 AVWWSPTGR 273

RESULT 35
Q61PC7 CAEBR
ID Q61PC7 CAEBR PRELIMINARY; PRT; 836 AA.
AC Q61PC7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG07598.
GN Name=CBG07598;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN NUCLEOTIDE SEQUENCE.
RP The C. briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAA01000032; CAB63236.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Hypothetical protein.
SQ SEQUENCE 836 AA; 95596 MW; F09E922E48C5DBCF CRC64;

Query Match 72.3%; Score 47; DB 2; Length 836;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPTGK 10
DB 271 AVWWSPTGR 279

RESULT 36
Q4FNA2_9RICK
ID Q4FNA2_9RICK PRELIMINARY; PRT; 240 AA.
AC Q4FNA2_9RICK
```

AC	Q42812;
DT	01-JUN-1998 (TrEMBLrel. 06, Created)
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Prolyl dipeptidyl peptidase precursor (EC 3.4.14.5).
GN	Name=dppIV;
OS	Aspergillus oryzae.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC	Eurotiates; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX	NCBI_TaxID=5062;
OX	[1]
RN	NUCLEOTIDE SEQUENCE.
RP	Doumas A., Van den Broek P., Affolter M., Monod M.;
RT	"Characterisation of the prolyl dipeptidyl peptidase IV activity; IEA.
RT	(dppIV) from the koji mould Aspergillus oryzae.";
RL	Submitted (Oct-1997) to the EMBL/GenBank/DDJB databases.
DR	EMBL; AJ002369; CAA05343.1; -; Genomic DNA.
DR	HSSP; P27487; LPFG.
DR	MEROPS; S09.008; -.
DR	GO; GO:0016020; C-membrane; IEA.
DR	GO; GO:0004274; F-dipeptidyl-peptidase IV activity; IEA.
DR	GO; GO:0016787; F-hydrolase activity; IEA.
DR	GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR	InterPro; IPR000425; MIP.
DR	InterPro; IPR001375; Peptidase_S9.
DR	InterPro; IPR002469; Peptidase_S9B.
DR	InterPro; IPR000379; Ser_estrs.
DR	Pfam; PF00930; DPPIV_N; 1.
DR	Pfam; PF00326; Peptidase_S9; 1.
DR	PROSITE; PS00221; MIP; UNKNOWN_1.
KW	Hydrolase; Signal.
FT	SIGNAL 1 16 Potential
FT	CHAIN 17 771 Prolyl dipeptidyl peptidase.
SQ	SEQUENCE 771 AA; 87208 MW; 8AB0CB33797B3FC6 CRC64;
Query Match	70.8%; Score 46; DB 2; Length 771;
Best Local Similarity	70.0%; Pred. No. 88;
Matches	7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY	1 YALWWSPNGK 10
DB	:: :
	205 YALWFSFDGE 214
RESULT 39	
Q7UKM8_RHOBA	
ID	Q7UKM8_RHOBA PRELIMINARY; PRT; 217 AA.
AC	Q7UKM8;
DT	01-OCT-2003 (TrEMBLrel. 25, Created)
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Probable alternative sigma factor Algu.
DN	OrderedLocustNames=RB10049;
OS	Rhodopirellula baltica.
OC	Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC	Planctomycetaceae; Firellula.
OX	NCBI_TaxID=117;
OX	[1]
RN	NUCLEOTIDE SEQUENCE.
RP	STRAIN=1;
RC	MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RX	Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA	Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA	Schleimer H., Amann R., Reinhardt R.;
RT	"Complete genome sequence of the marine planctomycete Firellula sp.
RT	strain 1,";
RL	Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RL	EMBL; BX294150; CAD76604.1; -; Genomic DNA.
DR	GO; GO:0003899; F-DNA-directed RNA polymerase activity; IEA.
DR	GO; GO:0016987; F-sigma factor activity; IEA.
DR	GO; GO:0003700; F-transcription factor activity; IEA.
DR	GO; GO:0016740; F-transferase activity; IEA.
DR	GO; GO:0006355; F-regulation of transcription, DNA-dependent; IEA.


```
DR GO:0006352; P:transcription initiation; IEA.
DR InterPro: IPR007627; Sigma70_r2.
DR Pfam: PF04542; Sigma70_r2; 1.
KW Complete proteome.
SQ SEQUENCE 217 AA; 25237 MW; E21F5D752873AFFA CRC64;

Query Match 69.2%; Score 45; DB 2; Length 217;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YALWSPNG 9
Db 12 YITWRRPG 20

RESULT 40
UBIG AGRTS STANDARD; PRT; 250 AA.
AC Q8UA66;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
GN Name=ubiG; OrderedLocusNames=Atu3508, AGR L.2640;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]_
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328 (2001).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -|- PATHWAY: Ubiquinone biosynthesis
CC -|- SIMILARITY: Belongs to the methyltransferase superfamily.
CC UbiG/COO3 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB009281; AAL44321.1; -; Genomic DNA.
CC EMBL; AB008332; AAK89887.1; -; Genomic_DNA.
CC PIR; AC2988; AC2988.

DR PIR; E98295; E98295.
DR HAMAP; MF 00472; -; 1.
DR InterPro: IPR003333; CMAS.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000051; SAM_bd.
DR InterPro: IPR010233; UbiG_mtfase.
DR Pfam; PF02353; CMAS; 1.
DR TIGRFAMs; TIGR01983; UbiG; 1.
KW Complete proteome; Methyltransferase; Transferase;
KW Ubiquinone biosynthesis.
SQ SEQUENCE 250 AA; 27827 MW; 2448F6627D093B88 CRC64;

Query Match 69.2%; Score 45; DB 1; Length 250;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWWSPPNGK 10
Db 22 AEWWSPTGK 30

RESULT 41
Q9XUK6 CAEL
ID Q9XUK6 CAEL PRELIMINARY; PRT; 320 AA.
AC Q9XUK6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein srxa-1.
GN Name=srxa-1; ORFNames=W07A8.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; Z82075; CAB04930.1; -; Genomic_DNA.
DR PIR; T26259; T26259.
DR Ensembl; W07A8.1; Caenorhabditis elegans.
DR WormBase; WBGene00012318; W07A8.1.
DR WormPep; W07A8.1; CE20159.
DR InterPro: IPR005047; DUF286; 1.
DR Pfam; PF03383; DUF286; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 320 AA; 36897 MW; 7B87D10281BB17D4 CRC64;

Query Match 69.2%; Score 45; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWSPPNG 9
Db 74 WWSPPNG 79

RESULT 42
Q7T290 BRABE
ID Q7T290 BRABE PRELIMINARY; PRT; 579 AA.
AC Q7T290;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein zgc:64214.
GN ORFNames=zgc:64214;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
```

OC Cyprinidae; Danilo.
OX NCBI_TaxID=7955;
RN [1] NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RX MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinaki M.I., Skaleka U., Snailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054643; AAH54643.1; -; mRNA.
DR HSSP; P38398; 1JM7.
DR Ensembl; ENSDARG0000007730; Danio rerio.
DR ZFIN; ZDB-GENE-040426-1407; zgc:64214.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001870; B302.
DR InterPro; IPR003879; Butyrophilin.
DR InterPro; IPR006574; PRY.
DR InterPro; IPR003877; SPRY receptor.
DR InterPro; IPR000315; Znf-Box.
DR InterPro; IPR001841; Znf-RING.
DR Pfam; PF00622; SPRY; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PRINTS; PR01406; BBOX2NFINGER.
DR PRINTS; PR01407; BUTYRPHLNCDUF.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00184; RING; 1.
DR SMART; SM00589; PRY; 1.
DR PROSITE; PS00119; ZF-BBOX; 1.
DR PROSITE; PS00518; ZF-RING; 1.
DR PROSITE; PS00899; ZF-RING_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 579 AA; 65428 MW; 24B23B304EC10C80 CRC64;

Query Match 69.2%; Score 45; DB 2; Length 579;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSFN 8
DB 497 YALWWSGN 504
|||||

RESULT 43
Q4LY82_9BURK
ID Q4LY82_9BURK PRELIMINARY; PRT; 787 AA.
AC Q4LY82;

DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Xanthine dehydrogenase (EC 1.1.1.7.1.4).
GN ORFNames=Bcen2424DRAFT_6085;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331272;
RN [1] NUCLEOTIDE SEQUENCE.
RP STRAIN=H12424;
RC US DOE Joint Genome Institute (JGI-PGF);
RG Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Istrati S., Pitluck S., Richardson P.,
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
H12424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2] NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.,
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
H12424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAHL01000005; EAM20974.1; -; Genomic_DNA.
KW Oxioreductase.
SQ SEQUENCE 787 AA; 84911 MW; 0AAC94973B3395AA CRC64;

Query Match 69.2%; Score 45; DB 2; Length 787;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWMSPNGK 10
DB 683 LWMNPGGK 690
|||||

RESULT 44
YDZF_SCHPO
ID YDZF_SCHPO STANDARD; PRT; 853 AA.
AC Q9P7E9; O13720;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Putative dipeptidyl aminopeptidase Cl4C4.15c (EC 3.4.14.-).
GN ORFNames=SPAC14C4.15c, SPAPU760.01c;
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1] NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Williams R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA Sgouras J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J.R., Voicckaert G., Aert R., Robben J., Grymonprez B.,

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Duesterhoeft A., Fritzc C., Holzer E., Moestl D.,
 RA Hilbert H., Borzyp K., Langer I., Beck A., Lehrach H., Reinhardt R.,
 RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Galleff A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Faulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Lysosome-like
 CC vacuoles.
 CC
 CC -|- VACUOLARITY: Belongs to the peptidase S9B family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AL162631; CAB83084.1; -; Genomic DNA.
 CC EMBL; Z98596; CAB11208.1; -; Genomic DNA.
 CC
 CC GenDB Spombe; SPAC14C4.15C; -;
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser esters.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC
 CC KW Aminopeptidase; Complete proteome; Glycoprotein; Hydrolase;
 KW Hypothetical protein; Protease; Serine protease; Signal-anchor;
 KW Transmembrane; Vacuole.
 FT TOPO_DOM 1 65
 FT TRANSMEM 66 86
 FT
 FT TOPO_DOM 87 853
 FT ACT_SITE 719 719
 FT ACT_SITE 795 795
 FT ACT_SITE 828 828
 FT CARBOHYD 96 96
 FT CARBOHYD 102 102
 FT CARBOHYD 472 472
 FT CARBOHYD 483 483
 FT CARBOHYD 613 613
 FT
 FT SEQUENCE 853 AA; 98342 MW; 38450BA50F8304B6 CRC64;
 SQ
 Query Match 69.2%; Score 45; DB 1; Length 853;
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ALWWSPPNG 9
 Db 288 AVWWSPDG 295
 | : | | | : |
 RESULT 45
 ID Q6PWQ4 XENLA PRELIMINARY; PRT; 227 AA.
 AC Q6PWQ4
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Matrix metalloproteinase 14 (Fragment).
 GN Name=MMP-14;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RX PubMed=15305302; DOI=10.1002/dvdy.20113;
 RA Harrison M., Abu-Elmagd M., Grocott T., Yates C., Gavrilovic J.,
 RA Wheeler G.N.;
 RT "Matrix metalloproteinase genes in Xenopus development.";
 RL Dev. Dyn. 231:214-220(2004).
 DR EMBL; AY573379; AAT00546.1; -; mRNA.
 DR InterPro; IPR000585; Hemoexin.
 DR Pfam; PF00045; Hemoexin; 3.
 DR SMART; SM00120; HX; 3.
 DR PROSITE; PS00024; HEMOEXIN; 1.
 FT NON_TER 1
 SQ SEQUENCE 227 AA; 25910 MW; 6A03B9B0ECF111FF CRC64;
 Query Match 67.7%; Score 44; DB 2; Length 227;
 Best Local Similarity 77.8%; Pred. No. 51;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ALWWSPPNGK 10
 Db 66 ALYWPNGK 74
 | : | | | |
 RESULT 46
 ID Q9ZN63 PRERU PRELIMINARY; PRT; 322 AA.
 AC Q9ZN63;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cellulase.
 OS Prevotella ruminicola (Bacteroides ruminicola).
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Prevotellaceae; Prevotella.
 OC NCBI_TaxID=839;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=JCM8958;
 RA Ogata K., Aminov R.I., Matsui H., Tajima K., Nakamura M., Nagamine T.,
 RA Benno Y.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB022867; BAA74515.1; -; Genomic DNA.
 DR HSSP; P07985; 1CEO.
 DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR001547; Glyco_hydro_5.
 DR Pfam; PF00150; Cellulase; 1.
 DR SEQUENCE 322 AA; 37331 MW; 413F0A69652F5237 CRC64;
 SQ
 Query Match 67.7%; Score 44; DB 2; Length 322;
 Best Local Similarity 60.0%; Pred. No. 73;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 YALWWSPPNGK 10
 Db 250 YGAWWTPLGK 259
 | : | | | : |
 RESULT 47
 ID Q7UP27 RHOB PRELIMINARY; PRT; 383 AA.
 AC Q7UP27;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=RH7205;
 OS Rhodopirella baltica.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Firellula.
 OC NCBI_TaxID=117;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=1;

```

RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schleener H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RL strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RL EMBL; BX294145; CAD75238.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 383 AA; 42610 MW; CAF119863344A439 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 383;
Best Local Similarity 72.7%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 2 ALWWSPP--NGK 10
Db 116 ALWWTDPEDNGK 126

RESULT 48
ID Q4QR7_LEIMA PRELIMINARY; PRT; 476 AA.
AC Q4QR7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE WD-40 repeat protein.
GN ORFNames=LmjF30.0410;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; CTO05267; CAJ05873.1; -; Genomic_DNA.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS00082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Antigen; Repeat; WD Repeat.
SQ SEQUENCE 476 AA; 5325 MW; 4A1389B4B60C57EF CRC64;

Query Match 67.7%; Score 44; DB 2; Length 476;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10
Db 135 YALQWSPSGR 144

RESULT 49
QSUSC3_XENLA
ID QSUSC3_XENLA PRELIMINARY; PRT; 575 AA.
AC QSUSC3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Membrane type-1 matrix metalloproteinase 14.
GN Name=MMP-14;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.

```

```

OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RT "Xenopus laevis MT1-MP.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY633953; AAU94940.1; -; mRNA.
DR SMR; QSUSC3; 105-275.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PRO0138; MATRXIN.
DR SMART; SM00120; HK; 4.
DR SMART; SM00235; ZmC; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Calcium; Hydrolase; Metal-binding; Metalloprotease; Protease; Zinc.
SQ SEQUENCE 575 AA; 65810 MW; EA8C4A9E9BB60D34 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 575;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
Db 414 ALYWPNGK 422

RESULT 50
Q6DCW8_XENLA
ID Q6DCW8_XENLA PRELIMINARY; PRT; 576 AA.
AC Q6DCW8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Mmp24-prov protein.
GN Name=mmp24-prov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RT TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC077870; AAH78780.1; -; mRNA.
DR SMR; Q6DCW8; 105-275.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR006026; Peptidase M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00045; Hemopexin_4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 576 AA; 65601 MW; F083F28DFD2C35F7 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 576;
Best Local Similarity 77.8%; Pred. No. 1.3e-02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
DB 415 ALYWPNGK 423
||| ||||
||| ||||

RESULT 51
Q5BL52 XENTR PRELIMINARY; PRT; 578 AA.
AC Q5BL52
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OC NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC090600; AAH90600.1; -; mRNA.
DR SMR; Q5BL52; 105-275.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR006026; Peptidase_M.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Calcium; Hydrolase; Hypothetical protein; Metal-binding;
KW Metalloprotease; Protease; Zinc.
SQ SEQUENCE 578 AA; 65838 MW; F7BC45C7EALFA96E CRC64;

Query Match 67.7%; Score 44; DB 2; Length 578;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
DB 417 ALYWPNGK 425
||| ||||
||| ||||

RESULT 52
MALQ HAEIN STANDARD; PRT; 699 AA.
AC P45176;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
DE (Disproportionation enzyme) (D-enzyme).
GN Name=malQ; OrderedLocustNames=H11356;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Priggs T., Hedblom E., Cotton M.D.,
RA Uterback T.T., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
 RL Science 269:496-512(1995).
 CC -|- CATALYTIC ACTIVITY: Transfers a segment of a 1,4-alpha-D-glucan to
 CC a new position in an acceptor, which may be glucose or a 1,4-
 CC alpha-D-glucan.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic. (By similarity).
 CC -|- SIMILARITY: Belongs to the dispropionating enzyme family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL: U32815; AAC23003.1; -; Genomic_DNA.
 DR PIR: H64118; H64118.
 DR TIGR: H11356; -.
 DR InterPro: IPR003385; Glyco_hydro_77.
 DR Pfam: PF02446; Glyco_hydro_77; 1.
 DR TIGRFAMs: TIGR00217; malQ_1.
 KW Carbohydrate metabolism; Complete proteome; Glycosyltransferase;
 KW Transferase.
 SQ SEQUENCE 699 AA; 80252 MW; 80D6E1D51EC2E1E9 CRC64;

Query Match 67.7%; Score 44; DB 1; Length 699;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
 : ||| |||
 Db 466 FRLWIPKPK 475

RESULT 53
 Q4QK66 HAE18 PRELIMINARY; PRT; 699 AA.
 AC Q4QK66;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE 4-alpha-glucanotransferase (EC 2.4.1.25).
 GN Name=malQ; OrderedLocNames=NT111810;
 OS Haemophilus influenzae (strain 86-028NP).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=281310;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
 RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,
 RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
 RA Munson R.S. Jr.;
 RT "Genomic sequence of an otitis media isolate of nontypeable
 RT Haemophilus influenzae: comparative study with H. influenzae serotype
 RT d, strain KW20.";
 RL J. Bacteriol. 187:4627-4636(2005).
 DR EMBL: CP000057; AAX88581.1; -; Genomic_DNA.
 DR InterPro: IPR003385; Glyco_hydro_77.
 DR Pfam: PF02446; Glyco_hydro_77; 1.
 DR TIGRFAMs: TIGR00217; malQ_1.
 KW Complete proteome; Glycosyltransferase; Transferase.
 SQ SEQUENCE 699 AA; 80147 MW; 32B8DDB72F06A0C6 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 699;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
 : ||| |||
 Db 466 FRLWIPKPK 475

RESULT 54
 Q82L10 STRAW PRELIMINARY; PRT; 1256 AA.
 ID Q82L10;
 AC Q82L10;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocNames=SAV2202;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt.820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL: BA000030; BAC69913.1; -; Genomic_DNA.
 DR GO: GO:0005488; F:binding; IEA.
 DR InterPro: IPR002114; HPR_Serp_S.
 DR InterPro: IPR011990; TPR-like_helical.
 DR PROSITE: PS00589; PTS_HPR_SER; UNKNOWN 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 1256 AA; 134840 MW; 73578B572F0F7C4C CRC64;

Query Match 67.7%; Score 44; DB 2; Length 1256;
 Best Local Similarity 85.7%; Pred. No. 3.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LWWSPNG 9
 : ||| |||
 Db 1002 LWWSPSG 1008

RESULT 55
 Q7Q7Z9 GIALA PRELIMINARY; PRT; 1271 AA.
 ID Q7Q7Z9;
 AC Q7Q7Z9;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE GLP_169_4270_455.
 OS Giardia lamblia ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 OX NCBI_TaxID=184922;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -|- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACB01000099; EAA38489.1; -; Genomic DNA.
 SQ SEQUENCE 1271 AA; 142786 MW; 712CA9D723B2490 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 1271;
 Best Local Similarity 85.7%; Pred. No. 3.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSP 7
 :|||||
 Db 386 FALWSP 392

RESULT 56
 Q9XYM9 TRIVA PRELIMINARY; PRT; 2320 AA.
 AC Q9XYM9 TRIVA PRELIMINARY; PRT; 2320 AA.
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Pre-mRNA processing 8 protein homolog PRP8.
 GN Name=PRP8;
 OS Trichomonas vaginalis.
 OC Eukaryota; Parabasalides; Trichomonada; Trichomonadida;
 OC Trichomonadidae; Trichomonadinae; Trichomonas.
 OX NCBI_TaxID=5722;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC30001;
 RX MEDLINE=99270317; PubMed=10340492; DOI=10.1016/S0166-6851(99)00017-1;
 RA Fast N.M., Doolittle W.F.;
 RT "Trichomonas vaginalis possesses a gene encoding the essential
 RT spliceosomal component, PRP8."
 RL Mol. Biochem. Parasitol. 99:275-278(1999).
 DR EMBL; AF115849; AAD29088.1; -; Genomic DNA.
 DR GO; GO:0008180; C:signalosome complex; IEA.
 DR InterPro; IPR000555; Mov34_MPN_PAD1.
 DR Pfam; PF01398; Mov34; 1.
 DR SMART; SM00232; JAB_MPN; 1.
 SQ SEQUENCE 2320 AA; 270718 MW; 5B3F4CF84DB25473 CRC64;

Query Match 67.7%; Score 44; DB 2; Length 2320;
 Best Local Similarity 85.7%; Pred. No. 5.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YALWSP 7
 :|||||
 Db 1524 YTLWSP 1530

RESULT 57
 Q4FW74 LEIMA PRELIMINARY; PRT; 2427 AA.
 AC Q4FW74 LEIMA PRELIMINARY; PRT; 2427 AA.
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE U5 snRNA-associated splicing factor.
 GN ORFNames=LWJ_1325;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RX PubMed=16020728; DOI=10.1126/science.1112680;
 RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
 RA Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
 RA Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
 RA Bianchetti G., Borzym K., Bothe G., Brusch C.V., Collins M.,
 RA Cadag E., Ciaroni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
 RA Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
 RA Fazlina G., Fosker N., Fraser A.C., Fraser A., Fuchs M., Gabel C.,
 RA Goble A., Goffeau A., Harris D., Hertz-Powder C., Hilbert H., Horn D.,
 RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
 RA Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,

RA Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
 RA Oliver K., O'Neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
 RA Quail M.A., Rabinowitz E., Reinhardt R., Rieger M., Rinta J.,
 RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D., Sharp S.,
 RA Schaffer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
 RA Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
 RA Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
 RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
 RA Myler P.J.;
 RT "The genome of the Kinetoplastid Parasite, Leishmania major.";
 RL Science 309:436-442(2005).
 DR EMBL; CP000081; AA214619.1; -; Genomic DNA.
 SQ SEQUENCE 2427 AA; 277492 MW; AAAD82891F4E5ECF CRC64;

Query Match 67.7%; Score 44; DB 2; Length 2427;
 Best Local Similarity 85.7%; Pred. No. 6.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSP 7
 :|||||
 Db 1608 FALWSP 1614

RESULT 58
 O97657 CANFA PRELIMINARY; PRT; 148 AA.
 ID O97657 CANFA PRELIMINARY; PRT; 148 AA.
 AC O97657 CANFA PRELIMINARY; PRT; 148 AA.
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Matrix metalloproteinase-14 (Fragment).
 GN Name=MMP-14;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxID=9615;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Carter S.D., Barnes A., Clegg P.D.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF097638; AAC64975.1; -; mRNA.
 DR HSP; P08253; 1CK7.
 DR MEROPS; M10.014; -.
 DR Ensembl; ENSCAFG00000011197; Canis familiaris.
 DR InterPro; IPR000585; Hemopexin.
 DR Pfam; PF00045; Hemopexin; 3.
 DR SMART; SM00120; HX; 3.
 FT NON_TER 1 148
 FT NON_TER 148 148
 SQ SEQUENCE 148 AA; 17366 MW; 2E5BC52426681A2B CRC64;

Query Match 66.2%; Score 43; DB 2; Length 148;
 Best Local Similarity 77.8%; Pred. No. 47;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWSPNGK 10
 :|:|:|:|
 Db 47 ALFWMPNGK 55

RESULT 59
 Q6QLX0 CANFA PRELIMINARY; PRT; 148 AA.
 ID Q6QLX0 CANFA PRELIMINARY; PRT; 148 AA.
 AC Q6QLX0 CANFA PRELIMINARY; PRT; 148 AA.
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Matrix metalloproteinase 14 (fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.

Query Match 66.2%; Score 43; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LWWSPN 8
DB 51 LWWSPN 56
Search completed: March 11, 2006, 12:24:53
Job time : 113.895 secs

NCBI_TaxID=9615;
[1]
NUCLEOTIDE SEQUENCE.
Groeters S., Alldinger S., Baumgaertner W.,
Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
EMBL; AY34615; AAS46101.1; -; mRNA.
DR Ensembl; ENSCAFG0000011197; Canis familiaris.
DR InterPro; IPR000585; Hemopexin.
DR Pfam; PF00045; Hemopexin; 3.
DR SMART; SM00120; HK; 3.
FT NON_TER 1 148
FT NON_TER 148 148
SQ SEQUENCE 148 AA; 17414 MW; EC12EB584E23289 CRC64;

Query Match 66.2%; Score 43; DB 2; Length 148;
Best Local Similarity 77.8%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 ALWWSPEK 10
DB 47 ALFWMEK 55

RESULT 60
LFTR LEGPA STANDARD; PRT; 222 AA.
AC OSX4EA;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Leucyl/phenylalanyl-tRNA--protein transferase (EC 2.3.2.6) (L/F-
transferase) (Leucyltransferase) (Phenylalanyltransferase).
GN Name-aat: OrderedLocusNames:lppl732;
OS Legionella pneumophila (strain Paris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=297246;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=15467720; DOI=10.1038/ng1447;
RX Cazalet C., Ruanio C., Brueggemann H., Zidane N., Magnier A., Ma L.,
Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
Etienne J., Glaser P., Buchrieser C.;
RT "Evidence in the Legionella pneumophila genome for exploitation of
host cell functions and high genome plasticity.";
RL Nat. Genet. 36:1165-1173(2004).
CC -!- FUNCTION: Functions in the N-end rule pathway of protein
degradation where it conjugates Leu, Phe and, less efficiently,
Met from aminoacyl-tRNAs to the N-termini of proteins containing
an N-terminal arginine or lysine (By similarity).
CC -!- CATALYTIC ACTIVITY: L-leucyl-tRNA + protein = tRNA + L-leucyl-
protein.
CC -!- CATALYTIC ACTIVITY: L-phenylalanyl-tRNA + protein = tRNA + L-
phenylalanyl-protein.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the L/F-transferase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
CC EMBL; CRG28336; CAH12884.1; -; Genomic_DNA.
CC Legiolist; lpp1732; -;
CC HAMAP; MF_00688; -; 1.
CC InterPro; IPR004616; Aat.
CC Pfam; PF03588; Leu_Phe_Trans; 1.
CC ProDom; PD022844; Aat; 1.
CC TIGRFAMs; TIGR00667; aat; 1.
CC Acyltransferase; Complete proteome; Transferase.
SQ SEQUENCE 222 AA; 25474 MW; BBB70BE39AF87BCE CRC64;

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:25:33 ; Search time 22.3684 Seconds
(without alignments)
36.961 Million cell updates/sec

Title: US-10-774-242A-4
Perfect score: 65
Sequence: 1 YALWSPNGK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	65	100.0	547	2	US-09-949-016-8330
2	65	100.0	760	1	US-08-230-491A-2
3	65	100.0	760	1	US-08-619-280A-2
4	65	100.0	760	1	US-08-940-391-2
5	65	100.0	760	2	US-09-265-606-2
6	60	92.3	761	2	US-09-518-550-26
7	53	81.5	593	4	PCT-US93-07923-11
8	53	81.5	755	4	PCT-US93-07923-3
9	53	81.5	759	4	PCT-US93-07923-2
10	53	81.5	766	1	US-08-230-491A-3
11	53	81.5	766	1	US-08-619-280A-3
12	53	81.5	766	1	US-08-940-391-3
13	53	81.5	766	2	US-09-794-236-1
14	53	81.5	766	2	US-10-002-593-6
15	53	81.5	766	2	US-09-949-016-6146
16	53	81.5	766	2	US-09-265-606-3
17	53	81.5	766	2	US-09-518-550-27
18	53	81.5	766	2	US-10-423-714-6
19	53	81.5	775	2	US-09-949-016-10450
20	46	70.8	771	2	US-09-462-284-2
21	46	70.8	771	2	US-09-079-592-2
22	45	69.2	574	2	US-09-902-540-11805
23	44	67.7	323	1	US-08-638-543-4
24	43	66.2	579	2	US-08-704-711A-1
25	43	66.2	579	2	US-09-521-220-1
26	43	66.2	582	2	US-08-704-711A-2
27	43	66.2	582	2	US-08-448-489-1

28	43	66.2	582	2	US-09-211-704A-9	Sequence 9, Appli
29	43	66.2	582	2	US-09-521-220-2	Sequence 2, Appli
30	43	66.2	582	2	US-09-391-104-28	Sequence 28, Appli
31	43	66.2	582	2	US-09-919-497-84	Sequence 84, Appli
32	43	66.2	582	2	US-09-689-730-1	Sequence 1, Appli
33	43	66.2	612	2	US-09-248-796A-18040	Sequence 18040, A
34	42	64.6	352	2	US-09-198-452A-702	Sequence 702, App
35	42	64.6	396	2	US-09-248-796A-1923	Sequence 1923, A
36	42	64.6	400	2	US-09-543-681A-5821	Sequence 5821, Ap
37	42	64.6	443	2	US-09-107-532A-7119	Sequence 7119, Ap
38	42	64.6	456	2	US-09-438-185A-667	Sequence 667, App
39	42	64.6	691	2	US-09-976-674-43	Sequence 43, Appli
40	42	64.6	706	2	US-09-976-674-41	Sequence 41, Appli
41	42	64.6	743	2	US-10-363-937-4	Sequence 4, Appli
42	42	64.6	796	2	US-09-976-674-5	Sequence 5, Appli
43	42	64.6	818	2	US-09-462-845-3	Sequence 3, Appli
44	42	64.6	818	2	US-10-402-312-3	Sequence 3, Appli
45	42	64.6	818	2	US-10-401-437-3	Sequence 3, Appli
46	42	64.6	818	2	US-10-402-067-3	Sequence 3, Appli
47	42	64.6	818	2	US-10-401-436-3	Sequence 3, Appli
48	41	63.1	320	2	US-09-248-796A-19283	Sequence 19283, A
49	41	63.1	564	2	US-09-211-704A-8	Sequence 8, Appli
50	41	63.1	669	2	US-08-704-711A-3	Sequence 3, Appli
51	41	63.1	669	2	US-09-521-220-3	Sequence 3, Appli
52	41	63.1	669	2	US-09-391-104-29	Sequence 29, Appli
53	40	61.5	178	2	US-09-902-540-10963	Sequence 10963, A
54	40	61.5	211	2	US-09-449-218D-12	Sequence 12, Appli
55	40	61.5	211	2	US-09-668-529A-12	Sequence 12, Appli
56	40	61.5	211	2	US-09-668-037A-12	Sequence 12, Appli
57	40	61.5	211	2	US-09-668-021-12	Sequence 12, Appli
58	40	61.5	213	2	US-09-449-218D-14	Sequence 14, Appli
59	40	61.5	213	2	US-09-668-529A-14	Sequence 14, Appli
60	40	61.5	213	2	US-09-668-037A-14	Sequence 14, Appli
61	40	61.5	213	2	US-09-668-021-14	Sequence 14, Appli
62	40	61.5	288	2	US-09-883-249-2	Sequence 2, Appli
63	40	61.5	351	2	US-09-248-796A-20975	Sequence 20975, A
64	40	61.5	421	2	US-10-067-443-4	Sequence 4, Appli
65	40	61.5	421	2	US-10-067-443-28	Sequence 28, Appli
66	40	61.5	482	2	US-09-270-767-58005	Sequence 58005, A
67	40	61.5	518	2	US-09-270-767-42683	Sequence 42683, A
68	40	61.5	545	2	US-09-270-767-58384	Sequence 58384, A
69	40	61.5	1088	2	US-09-270-767-43052	Sequence 43052, A
70	39	60.0	2357	2	US-09-949-016-9656	Sequence 9656, Ap
71	39	60.0	75	2	US-09-252-991A-21149	Sequence 21149, A
72	39	60.0	174	2	US-09-914-098-16	Sequence 16, Appli
73	39	60.0	253	2	US-09-270-767-34417	Sequence 34417, A
74	39	60.0	253	2	US-09-270-767-49634	Sequence 49634, A
75	39	60.0	255	2	US-09-914-098-42	Sequence 42, Appli
76	39	60.0	766	2	US-09-487-558B-316	Sequence 316, App
77	39	60.0	981	2	US-09-902-540-16812	Sequence 16812, A
78	39	60.0	1070	2	US-09-107-532A-3716	Sequence 3716, Ap
79	39	60.0	1073	2	US-09-134-000C-6356	Sequence 6356, Ap
80	39.5	59.2	207	2	US-09-214-909-24	Sequence 24, Appli
81	38	58.5	104	2	US-10-104-047-2562	Sequence 2562, Ap
82	38	58.5	127	2	US-09-270-767-38926	Sequence 38926, A
83	38	58.5	127	2	US-09-270-767-54143	Sequence 54143, A
84	38	58.5	210	2	US-09-252-991A-23899	Sequence 23899, A
85	38	58.5	237	2	US-09-543-681A-5550	Sequence 5550, Ap
86	38	58.5	242	2	US-09-489-039A-12133	Sequence 12133, A
87	38	58.5	332	2	US-09-252-991A-31442	Sequence 31442, A
88	38	58.5	355	2	US-09-733-524A-12	Sequence 12, Appli
89	38	58.5	355	2	US-10-189-977A-12	Sequence 12, Appli
90	38	58.5	356	2	US-09-092-315-12	Sequence 12, Appli
91	38	58.5	367	2	US-09-252-991A-16628	Sequence 16628, A
92	38	58.5	393	2	US-09-390-131-8	Sequence 8, Appli
93	38	58.5	413	2	US-09-328-352-5589	Sequence 5589, Ap
94	38	58.5	451	2	US-09-252-991A-25870	Sequence 25870, A
95	37	56.9	157	2	US-09-489-039A-9109	Sequence 9109, Ap
96	37	56.9	220	2	US-09-252-991A-22823	Sequence 22823, A
97	37	56.9	242	2	US-09-248-796A-20477	Sequence 20477, A
98	37	56.9	258	1	US-08-336-708A-10	Sequence 10, Appli
99	37	56.9	438	2	US-09-339-838-5	Sequence 5, Appli
100	37	56.9	438	2	US-09-339-838-7	Sequence 7, Appli

101	37	56.9	508	1	US-08-850-293-5	Sequence 5, Appli	174	35	53.8	71	2	US-09-370-767-44774	Sequence 44774, A
102	37	56.9	533	2	US-08-949-016-9753	Sequence 9753, Ap	175	35	53.8	87	2	US-09-732-210-359	Sequence 359, App
103	37	56.9	567	2	US-09-489-039A-10840	Sequence 10840, A	176	35	53.8	112	2	US-09-489-039A-10236	Sequence 10236, A
104	37	56.9	698	2	US-09-248-796A-19233	Sequence 19233, A	177	35	53.8	134	2	US-09-673-395A-356	Sequence 356, App
105	37	56.9	762	2	US-09-228-986-114	Sequence 114, App	178	35	53.8	162	2	US-09-252-991A-17568	Sequence 17568, A
106	37	56.9	762	2	US-10-101-464A-114	Sequence 114, App	179	35	53.8	175	2	US-09-252-991A-17111	Sequence 17111, A
107	37	56.9	766	2	US-09-248-796A-16626	Sequence 16626, A	180	35	53.8	176	2	US-09-449-218D-16	Sequence 16, Appl
108	37	56.9	931	2	US-09-079-592-11	Sequence 11, Appl	181	35	53.8	176	2	US-09-668-529A-16	Sequence 16, Appl
109	37	56.9	1178	2	US-09-438-185A-851	Sequence 851, App	182	35	53.8	176	2	US-09-668-037A-16	Sequence 16, Appl
110	36	55.4	91	2	US-09-513-999C-4160	Sequence 4160, Ap	183	35	53.8	176	2	US-09-668-021-16	Sequence 16, Appl
111	36	55.4	93	1	US-08-447-464-5	Sequence 5, Appli	184	35	53.8	180	2	US-09-902-540-14375	Sequence 2, Appli
112	36	55.4	93	1	US-08-716-679-5	Sequence 5, Appli	185	35	53.8	213	2	US-09-449-218D-6	Sequence 6, Appli
113	36	55.4	106	2	US-09-543-581A-5145	Sequence 5145, Ap	186	35	53.8	213	2	US-09-449-218D-8	Sequence 8, Appli
114	36	55.4	131	2	US-08-981-392-60	Sequence 60, Appl	187	35	53.8	213	2	US-09-449-218D-10	Sequence 10, Appli
115	36	55.4	131	2	US-09-908-322-60	Sequence 60, Appl	188	35	53.8	213	2	US-09-668-529A-2	Sequence 2, Appli
116	36	55.4	146	2	US-09-270-767-36512	Sequence 36512, A	189	35	53.8	213	2	US-09-668-529A-6	Sequence 6, Appli
117	36	55.4	146	2	US-09-270-767-51729	Sequence 51729, A	190	35	53.8	213	2	US-09-668-529A-8	Sequence 8, Appli
118	36	55.4	181	2	US-09-538-092-792	Sequence 792, App	191	35	53.8	213	2	US-09-668-037A-10	Sequence 10, Appl
119	36	55.4	245	1	US-08-321-670-3	Sequence 3, Appli	192	35	53.8	213	2	US-09-668-037A-6	Sequence 6, Appli
120	36	55.4	249	2	US-09-270-767-32523	Sequence 32523, A	193	35	53.8	213	2	US-09-668-037A-8	Sequence 8, Appli
121	36	55.4	249	2	US-09-270-767-47740	Sequence 47740, A	194	35	53.8	213	2	US-09-668-037A-10	Sequence 10, Appl
122	36	55.4	290	2	US-09-489-039A-14080	Sequence 14080, A	195	35	53.8	213	2	US-09-668-037A-10	Sequence 10, Appl
123	36	55.4	357	4	PCT-US91-00899-14	Sequence 14, Appl	196	35	53.8	213	2	US-09-668-021-2	Sequence 2, Appli
124	36	55.4	393	2	US-09-303-518D-876	Sequence 876, App	197	35	53.8	213	2	US-09-668-021-6	Sequence 6, Appli
125	36	55.4	400	2	US-09-270-767-31765	Sequence 31765, A	198	35	53.8	213	2	US-09-668-021-8	Sequence 8, Appli
126	36	55.4	400	2	US-09-270-767-46982	Sequence 46982, A	199	35	53.8	213	2	US-09-668-021-10	Sequence 10, Appl
127	36	55.4	405	1	US-07-914-281-8	Sequence 8, Appli	200	35	53.8	213	2	US-09-605-703B-2342	Sequence 3, Appli
128	36	55.4	405	1	US-08-393-246-8	Sequence 8, Appli	201	35	53.8	237	2	US-08-295-657-3	Sequence 3, Appli
129	36	55.4	405	1	US-08-525-058A-8	Sequence 8, Appli	202	35	53.8	295	1	US-09-579-664B-12	Sequence 12, Appl
130	36	55.4	405	1	US-08-483-151-4	Sequence 4, Appli	203	35	53.8	311	2	US-10-355-975A-12	Sequence 12, Appl
131	36	55.4	405	1	US-08-696-731-8	Sequence 8, Appli	204	35	53.8	311	2	US-09-823-038A-59	Sequence 59, Appl
132	36	55.4	405	2	US-09-042-531-8	Sequence 8, Appli	205	35	53.8	311	2	US-08-362-739-2	Sequence 2, Appli
133	36	55.4	426	2	US-09-602-787A-424	Sequence 424, App	206	35	53.8	320	1	US-08-914-350A-2	Sequence 2, Appli
134	36	55.4	464	2	US-08-252-991A-24883	Sequence 24883, A	207	35	53.8	320	2	US-09-489-039A-11458	Sequence 11458, A
135	36	55.4	533	1	US-08-294-872-2	Sequence 2, Appli	208	35	53.8	344	2	US-08-447-702-5	Sequence 5, Appli
136	36	55.4	533	2	PCT-US95-09823-2	Sequence 2, Appli	209	35	53.8	354	1	US-08-465-615-5	Sequence 5, Appli
137	36	55.4	584	2	US-09-595-857B-25	Sequence 25, Appl	210	35	53.8	354	2	US-09-179-558-65	Sequence 65, Appl
138	36	55.4	588	2	US-09-949-016-10547	Sequence 10547, A	211	35	53.8	359	2	US-09-722-825-65	Sequence 65, Appl
139	36	55.4	743	2	US-09-489-039A-11073	Sequence 11073, A	212	35	53.8	359	2	US-09-722-825-65	Sequence 65, Appl
140	36	55.4	848	2	US-09-787-443-44	Sequence 44, Appl	213	35	53.8	359	2	US-09-722-708-65	Sequence 65, Appl
141	36	55.4	871	2	US-09-303-518D-880	Sequence 880, App	214	35	53.8	359	2	US-09-489-039A-9541	Sequence 9541, Ap
142	36	55.4	888	2	US-09-303-518D-878	Sequence 878, App	215	35	53.8	398	2	US-09-489-039A-12259	Sequence 12259, A
143	36	55.4	922	2	US-09-303-518D-882	Sequence 882, App	216	35	53.8	429	2	US-08-576-626A-54	Sequence 54, Appl
144	36	55.4	922	2	US-09-303-518D-884	Sequence 884, App	217	35	53.8	431	1	US-08-576-626A-54	Sequence 54, Appl
145	36	55.4	1290	1	US-08-470-350B-2	Sequence 2, Appli	218	35	53.8	431	2	US-09-522-991A-23720	Sequence 23720, A
146	36	55.4	3559	2	US-09-693-205A-10	Sequence 10, Appl	219	35	53.8	472	2	US-09-489-039A-18086	Sequence 18086, A
147	35.5	54.6	178	2	US-09-903-456-50	Sequence 50, Appl	220	35	53.8	473	2	US-09-489-039A-9854	Sequence 9854, Ap
148	35.5	54.6	178	2	US-09-624-670-49	Sequence 49, Appl	221	35	53.8	481	2	US-09-248-796A-19316	Sequence 19316, A
149	35.5	54.6	182	2	US-09-903-456-46	Sequence 46, Appl	222	35	53.8	484	2	US-10-104-047-3591	Sequence 3591, Ap
150	35.5	54.6	182	2	US-09-624-670-45	Sequence 45, Appl	223	35	53.8	497	2	US-09-489-039A-12027	Sequence 12027, A
151	35.5	54.6	238	2	US-09-903-456-42	Sequence 42, Appl	224	35	53.8	506	2	US-09-489-039A-13073	Sequence 13073, A
152	35.5	54.6	238	2	US-09-624-670-41	Sequence 41, Appl	225	35	53.8	604	2	US-09-391-104-30	Sequence 30, Appl
153	35.5	54.6	241	2	US-09-903-456-48	Sequence 48, Appl	226	35	53.8	607	2	US-09-000-041A-2	Sequence 2, Appli
154	35.5	54.6	241	2	US-09-624-670-47	Sequence 47, Appl	227	35	53.8	607	2	US-09-211-704A-10	Sequence 10, Appl
155	35.5	54.6	265	2	US-09-903-456-54	Sequence 54, Appl	228	35	53.8	607	2	US-09-734-002-2	Sequence 2, Appli
156	35.5	54.6	265	2	US-09-624-670-53	Sequence 53, Appl	229	35	53.8	625	2	US-09-902-540-15330	Sequence 15330, A
157	35.5	54.6	265	2	US-09-903-456-44	Sequence 44, Appl	230	35	53.8	631	2	US-09-311-626B-16	Sequence 16, Appl
158	35.5	54.6	278	2	US-09-624-670-43	Sequence 43, Appl	231	35	53.8	649	2	US-09-248-796A-20641	Sequence 20641, A
159	35.5	54.6	278	2	US-09-903-456-32	Sequence 32, Appl	232	35	53.8	727	2	US-09-179-558-56	Sequence 56, Appl
160	35.5	54.6	278	2	US-09-624-670-31	Sequence 31, Appl	233	35	53.8	727	2	US-09-722-825-55	Sequence 55, Appl
161	35.5	54.6	280	2	US-09-903-456-87	Sequence 87, Appl	234	35	53.8	727	2	US-09-722-487-56	Sequence 56, Appl
162	35.5	54.6	282	2	US-09-903-456-57	Sequence 57, Appl	235	35	53.8	727	2	US-09-722-708-56	Sequence 56, Appl
163	35.5	54.6	302	2	US-09-624-670-56	Sequence 56, Appl	236	35	53.8	791	2	US-10-104-047-3453	Sequence 3453, Ap
164	35.5	54.6	302	2	US-09-903-456-52	Sequence 52, Appl	237	35	53.8	941	2	US-09-179-558-55	Sequence 55, Appl
165	35.5	54.6	302	2	US-09-624-670-51	Sequence 51, Appl	238	35	53.8	941	2	US-09-722-825-55	Sequence 55, Appl
166	35.5	54.6	306	2	US-09-605-703B-2850	Sequence 2850, Ap	239	35	53.8	941	2	US-09-722-708-55	Sequence 55, Appl
167	35.5	54.6	318	2	US-09-903-456-31	Sequence 31, Appl	240	35	53.8	941	2	US-09-289-349-10	Sequence 10, Appl
168	35.5	54.6	318	2	US-09-624-670-30	Sequence 30, Appl	241	34.5	53.1	139	2	US-09-621-976-4554	Sequence 4554, Ap
169	35.5	54.6	628	1	US-08-261-822A-6	Sequence 6, Appli	242	34.5	53.1	147	2	US-09-924-330A-9	Sequence 9, Appli
170	35.5	54.6	628	4	PCT-US95-07744A-6	Sequence 7, Appli	243	34	52.3	36	2	US-08-924-330A-9	Sequence 9, Appli
171	35	53.8	60	1	US-08-362-739-7	Sequence 7, Appli	244	34	52.3	40	2	US-09-138-721-9	Sequence 23, Appli
172	35	53.8	60	2	US-08-914-350A-7	Sequence 7, Appli	245	34	52.3	40	2	US-09-275-352A-23	Sequence 23, Appli
173	35	53.8	70	2	US-09-270-767-60232	Sequence 60232, A	246	34	52.3	40	2		

247	34	52.3	53	2	US-09-270-767-58810	Sequence 58810, A	320	34	52.3	548	2	US-09-543-681A-6920	Sequence 6920, Ap
248	34	52.3	54	2	US-09-621-976-7003	Sequence 7003, Ap	321	34	52.3	549	1	US-08-791-887-2	Sequence 2, Appl
249	34	52.3	58	2	US-09-275-252A-3	Sequence 3, Appl	322	34	52.3	549	2	US-09-146-084-2	Sequence 2, Appl
250	34	52.3	58	2	US-09-275-252A-25	Sequence 25, Appl	323	34	52.3	566	2	US-09-902-540-10183	Sequence 10183, A
251	34	52.3	59	2	US-09-384-302A-16	Sequence 16, Appl	324	34	52.3	608	2	US-08-843-572E-2	Sequence 2, Appl
252	34	52.3	62	2	US-09-513-999C-7437	Sequence 7437, Ap	325	34	52.3	627	2	US-09-345-473E-46	Sequence 46, Appl
253	34	52.3	62	1	US-08-344-695-15	Sequence 15, Appl	326	34	52.3	627	2	US-09-862-027-46	Sequence 46, Appl
254	34	52.3	74	2	US-09-270-767-35267	Sequence 35267, A	327	34	52.3	655	2	US-10-152-886-89	Sequence 89, Appl
255	34	52.3	72	2	US-09-270-767-50484	Sequence 50484, A	328	34	52.3	803	2	US-09-543-681A-7287	Sequence 7287, Ap
256	34	52.3	72	2	US-09-270-767-50484	Sequence 50484, A	329	34	52.3	855	2	US-09-543-681A-7287	Sequence 7287, Ap
257	34	52.3	72	2	US-09-270-767-50484	Sequence 50484, A	330	34	52.3	855	2	US-09-543-681A-7287	Sequence 7287, Ap
258	34	52.3	90	2	US-09-384-302A-17	Sequence 17, Appl	331	34	52.3	959	2	US-09-107-433-433A	Sequence 433A, Ap
259	34	52.3	92	2	US-09-902-540-10408	Sequence 10408, A	332	34	52.3	1032	2	US-09-489-039A-14025	Sequence 14025, A
260	34	52.3	112	2	US-09-252-991A-26083	Sequence 26083, A	333	34	52.3	1032	2	US-09-583-110-3366	Sequence 3366, Ap
261	34	52.3	145	2	US-09-252-991A-22680	Sequence 22680, A	334	34	52.3	1089	2	US-09-248-796A-19072	Sequence 19072, A
262	34	52.3	145	2	US-09-252-991A-18802	Sequence 18802, A	335	34	52.3	1248	2	US-10-042-810-2	Sequence 2, Appl
263	34	52.3	158	2	US-09-252-991A-27898	Sequence 27898, A	336	34	52.3	1278	2	US-10-042-810-2	Sequence 2, Appl
264	34	52.3	160	2	US-09-275-252A-1	Sequence 1, Appl	337	34	52.3	1338	2	US-09-631-603-4	Sequence 2, Appl
265	34	52.3	160	2	US-09-275-252A-16	Sequence 16, Appl	338	34	52.3	1338	2	US-09-631-603-4	Sequence 2, Appl
266	34	52.3	181	2	US-09-927-112-14	Sequence 14, Appl	339	34	52.3	1587	2	US-09-845-583A-10	Sequence 10, Appl
267	34	52.3	181	2	US-10-104-047-2677	Sequence 2677, Ap	340	34	52.3	1587	2	US-09-561-709B-3	Sequence 3, Appl
268	34	52.3	182	2	US-09-252-991A-29639	Sequence 29639, A	341	34	52.3	2972	2	US-08-469-260A-387	Sequence 387, App
269	34	52.3	182	2	US-09-583-110-4684	Sequence 4684, Ap	342	34	52.3	2972	2	US-08-469-260A-387	Sequence 387, App
270	34	52.3	184	2	US-09-543-681A-4219	Sequence 4219, Ap	343	34	52.3	2972	2	US-08-469-260A-387	Sequence 387, App
271	34	52.3	190	2	US-09-270-767-31656	Sequence 31656, A	344	34	52.3	2972	2	US-08-424-550B-387	Sequence 5109, Ap
272	34	52.3	192	2	US-09-570-858B-23	Sequence 9964, Ap	345	34	52.3	299	2	US-09-107-582A-5109	Sequence 89, Appl
273	34	52.3	194	2	US-09-570-858B-24	Sequence 24, Appl	346	34	52.3	12	1	US-08-190-788A-89	Sequence 89, Appl
274	34	52.3	194	2	US-09-570-858B-24	Sequence 24, Appl	347	34	52.3	12	1	US-08-383-474B-94	Sequence 94, Appl
275	34	52.3	197	2	US-09-248-796A-17262	Sequence 17262, A	348	34	52.3	12	1	US-08-465-391A-89	Sequence 89, Appl
276	34	52.3	217	2	US-09-270-767-33317	Sequence 33317, A	349	34	52.3	12	1	US-08-465-391A-89	Sequence 89, Appl
277	34	52.3	218	2	US-09-270-767-34355	Sequence 34355, A	350	34	52.3	12	2	US-08-465-391A-89	Sequence 89, Appl
278	34	52.3	225	1	US-08-886-765-2	Sequence 2, Appl	351	34	52.3	12	2	US-08-465-391A-89	Sequence 89, Appl
279	34	52.3	225	2	US-09-115-660-2	Sequence 2, Appl	352	34	52.3	12	2	US-08-465-391A-89	Sequence 89, Appl
280	34	52.3	225	2	US-09-902-540-15456	Sequence 15456, A	353	34	52.3	27	1	US-09-072-323-4	Sequence 4, Appl
281	34	52.3	228	2	US-09-893-737-142	Sequence 142, App	354	34	52.3	28	1	US-09-072-323-4	Sequence 6, Appl
282	34	52.3	229	2	US-09-252-991A-32410	Sequence 32410, A	355	34	52.3	32	1	US-08-190-802A-167	Sequence 167, App
283	34	52.3	234	2	US-09-902-540-14926	Sequence 14926, A	356	34	52.3	32	1	US-08-190-802A-212	Sequence 212, App
284	34	52.3	235	2	US-09-949-016-7903	Sequence 7903, Ap	357	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
285	34	52.3	245	2	US-08-469-260A-40	Sequence 40, Appl	358	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
286	34	52.3	245	2	US-08-469-260A-40	Sequence 40, Appl	359	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
287	34	52.3	245	2	US-08-469-260A-40	Sequence 40, Appl	360	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
288	34	52.3	245	2	US-08-469-260A-40	Sequence 40, Appl	361	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
289	34	52.3	256	2	US-09-543-681A-6189	Sequence 6189, Ap	362	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
290	34	52.3	266	2	US-09-270-767-44050	Sequence 44050, A	363	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
291	34	52.3	266	2	US-09-248-796A-17346	Sequence 17346, A	364	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
292	34	52.3	280	2	US-09-902-540-10405	Sequence 10405, A	365	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
293	34	52.3	285	1	US-08-712-072C-5	Sequence 5, Appl	366	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
294	34	52.3	294	2	US-10-087-013-7	Sequence 7, Appl	367	34	52.3	32	2	US-08-477-346-167	Sequence 167, App
295	34	52.3	314	2	US-09-252-991A-19600	Sequence 19600, A	368	34	52.3	52	2	US-08-894-017-7	Sequence 7, Appl
296	34	52.3	322	2	US-09-252-991A-31608	Sequence 31608, A	369	34	52.3	52	2	US-09-456-474-7	Sequence 40769, A
297	34	52.3	345	1	US-08-458-970A-2	Sequence 2, Appl	370	34	52.3	59	2	US-09-270-767-55985	Sequence 55985, A
298	34	52.3	355	1	US-08-458-970A-2	Sequence 2, Appl	371	34	52.3	59	2	US-09-270-767-55985	Sequence 55985, A
299	34	52.3	355	2	US-09-248-796A-17341	Sequence 17341, A	372	34	52.3	66	2	US-09-489-039A-11221	Sequence 11221, A
300	34	52.3	366	2	US-09-248-796A-17341	Sequence 17341, A	373	34	52.3	66	2	US-09-489-039A-11221	Sequence 11221, A
301	34	52.3	392	2	US-09-902-540-10171	Sequence 10171, A	374	34	52.3	68	2	US-09-227-357-277	Sequence 277, App
302	34	52.3	394	2	US-09-949-016-9551	Sequence 9551, Ap	375	34	52.3	68	2	US-09-227-357-277	Sequence 277, App
303	34	52.3	402	2	US-09-328-352-4953	Sequence 4953, Ap	376	34	52.3	70	1	US-08-315-695-3	Sequence 3, Appl
304	34	52.3	422	1	US-08-712-072C-2	Sequence 2, Appl	377	34	52.3	70	1	US-08-315-695-3	Sequence 3, Appl
305	34	52.3	427	2	US-09-252-991A-23397	Sequence 23397, A	378	34	52.3	71	2	US-09-252-991A-17632	Sequence 17632, A
306	34	52.3	431	2	US-09-252-991A-29948	Sequence 29948, A	379	34	52.3	72	2	US-09-248-796A-21385	Sequence 21385, A
307	34	52.3	431	2	US-09-198-452A-831	Sequence 831, App	380	34	52.3	74	2	US-09-227-357-163	Sequence 163, App
308	34	52.3	435	2	US-09-438-185A-784	Sequence 784, App	381	34	52.3	74	2	US-09-227-357-163	Sequence 206, App
309	34	52.3	440	2	US-09-248-796A-17146	Sequence 17146, A	382	34	52.3	74	2	US-09-227-357-163	Sequence 25368, A
310	34	52.3	445	2	US-10-104-047-2209	Sequence 2209, Ap	383	34	52.3	80	2	US-09-973-278-206	Sequence 25368, A
311	34	52.3	450	1	US-08-818-514-6	Sequence 6, Appl	384	34	52.3	80	2	US-09-973-278-206	Sequence 11965, A
312	34	52.3	450	2	US-09-115-934A-6	Sequence 6, Appl	385	34	52.3	91	2	US-09-489-039A-11965	Sequence 31, Appl
313	34	52.3	450	2	US-09-611-175-6	Sequence 6, Appl	386	34	52.3	94	2	US-09-830-807-31	Sequence 11, Appl
314	34	52.3	460	2	US-09-489-039A-11840	Sequence 11840, A	387	34	52.3	96	1	US-08-341-843B-11	Sequence 16, Appl
315	34	52.3	470	2	US-09-902-540-16052	Sequence 16052, A	388	34	52.3	100	1	US-08-427-497E-16	Sequence 22, Appl
316	34	52.3	472	2	US-09-605-703B-1400	Sequence 1400, Ap	389	34	52.3	100	2	US-07-736-335E-22	Sequence 22, Appl
317	34	52.3	472	2	US-09-605-703B-1426	Sequence 1426, Ap	390	34	52.3	101	2	US-09-270-767-35993	Sequence 35993, A
318	34	52.3	472	2	US-09-605-703B-1428	Sequence 1428, Ap	391	34	52.3	101	2	US-09-270-767-35993	Sequence 51210, A
319	34	52.3	532	2	US-09-489-039A-9142	Sequence 9142, Ap	392	34	52.3	101	2	US-09-270-767-35993	Sequence 24559, A

393	33	50.8	144	2	US-09-248-796A-26309	Sequence 26309, A	466	33	50.8	415	2	US-09-538-092-708	Sequence 708, App
394	33	50.8	150	2	US-09-252-991A-29818	Sequence 29818, A	467	33	50.8	425	1	US-08-190-802A-58	Sequence 58, Appl
395	33	50.8	154	2	US-09-252-991A-28772	Sequence 28772, A	468	33	50.8	425	2	US-08-105-454-2	Sequence 2, Appl
396	33	50.8	156	2	US-09-107-433-4460	Sequence 4460, Ap	469	33	50.8	425	2	US-08-477-346-58	Sequence 58, Appl
397	33	50.8	167	2	US-09-134-000C-6568	Sequence 6568, Ap	470	33	50.8	425	2	US-08-473-089-58	Sequence 58, Appl
398	33	50.8	167	2	US-09-134-000C-6693	Sequence 6693, Ap	471	33	50.8	425	2	US-08-487-072A-58	Sequence 17, Appl
399	33	50.8	167	2	US-09-949-016-7608	Sequence 7608, Ap	472	33	50.8	425	2	US-09-538-092-1288	Sequence 1288, Ap
400	33	50.8	170	2	US-09-198-452A-897	Sequence 897, App	473	33	50.8	425	2	US-08-624-735E-17	Sequence 17, Appl
401	33	50.8	172	2	US-09-252-991A-23050	Sequence 23050, A	474	33	50.8	425	2	US-09-949-016-6643	Sequence 6643, Ap
402	33	50.8	184	2	US-09-270-767-45190	Sequence 45190, A	475	33	50.8	426	2	US-09-252-991A-25192	Sequence 25192, A
403	33	50.8	193	1	US-08-464-342-6	Sequence 6, Appl	476	33	50.8	426	2	US-09-673-395A-134	Sequence 134, App
404	33	50.8	193	1	US-08-875-272-6	Sequence 6, Appl	477	33	50.8	432	2	US-09-443-041A-20	Sequence 20, Appl
405	33	50.8	193	1	US-08-903-396-6	Sequence 6, Appl	478	33	50.8	432	2	US-09-092-315-11	Sequence 11, Appl
406	33	50.8	194	2	US-08-679-493A-215	Sequence 215, App	479	33	50.8	433	2	US-09-733-524A-11	Sequence 11, Appl
407	33	50.8	194	2	US-08-679-493A-216	Sequence 216, App	480	33	50.8	433	2	US-10-189-977A-11	Sequence 11, Appl
408	33	50.8	202	2	US-09-489-039A-11773	Sequence 11773, A	481	33	50.8	434	2	US-09-489-039A-13162	Sequence 13162, A
409	33	50.8	204	2	US-09-248-796A-17274	Sequence 17274, A	482	33	50.8	452	1	US-08-984-171-3	Sequence 3, Appl
410	33	50.8	209	2	US-09-252-991A-19957	Sequence 19957, A	483	33	50.8	467	2	US-09-438-185A-349	Sequence 349, App
411	33	50.8	213	2	US-09-902-540-12104	Sequence 12104, A	484	33	50.8	477	2	US-09-252-991A-24574	Sequence 24574, A
412	33	50.8	220	2	US-09-252-991A-20618	Sequence 20618, A	485	33	50.8	488	2	US-09-252-991A-28535	Sequence 28535, A
413	33	50.8	228	2	US-09-286-690-12	Sequence 12, Appl	486	33	50.8	503	2	US-09-107-532A-6896	Sequence 6896, Ap
414	33	50.8	236	2	US-09-902-540-15416	Sequence 15416, A	487	33	50.8	512	2	US-09-902-540-10238	Sequence 10238, Ap
415	33	50.8	239	2	US-09-489-039A-13949	Sequence 13949, A	488	33	50.8	526	2	US-09-605-703B-1776	Sequence 1776, Ap
416	33	50.8	236	2	US-09-902-540-9715	Sequence 9715, Ap	489	33	50.8	540	2	US-09-963-668B-2	Sequence 2, Appl
417	33	50.8	245	2	US-09-303-518D-700	Sequence 700, App	490	33	50.8	546	2	US-09-907-794A-250	Sequence 250, App
418	33	50.8	245	2	US-09-303-518D-702	Sequence 702, App	491	33	50.8	546	2	US-09-905-125A-250	Sequence 250, App
419	33	50.8	245	2	US-09-303-518D-704	Sequence 704, App	492	33	50.8	546	2	US-09-902-775A-250	Sequence 250, App
420	33	50.8	247	2	US-09-228-986-105	Sequence 105, App	493	33	50.8	546	2	US-09-906-700-250	Sequence 250, App
421	33	50.8	247	2	US-10-101-464A-105	Sequence 105, App	494	33	50.8	546	2	US-09-903-603A-250	Sequence 250, App
422	33	50.8	250	2	US-09-328-352-5728	Sequence 5728, Ap	495	33	50.8	546	2	US-09-904-920A-250	Sequence 250, App
423	33	50.8	254	2	US-09-540-236-3796	Sequence 3796, Ap	496	33	50.8	546	2	US-09-909-064-250	Sequence 250, App
424	33	50.8	261	2	US-09-248-796A-23710	Sequence 23710, A	497	33	50.8	546	2	US-09-905-381A-250	Sequence 250, App
425	33	50.8	288	2	US-09-252-991A-25627	Sequence 25627, A	498	33	50.8	546	2	US-09-906-618-250	Sequence 250, App
426	33	50.8	300	2	US-09-949-016-11599	Sequence 11599, A	499	33	50.8	546	2	US-09-906-646-250	Sequence 250, App
427	33	50.8	303	2	US-09-420-786A-3	Sequence 3, Appl	500	33	50.8	546	2	US-09-904-462-250	Sequence 250, App
428	33	50.8	308	2	US-09-248-796A-19874	Sequence 19874, A							
429	33	50.8	311	1	US-08-568-459A-21	Sequence 21, Appl							
430	33	50.8	311	1	US-08-487-826B-33	Sequence 33, Appl							
431	33	50.8	311	2	US-09-210-288-21	Sequence 21, Appl							
432	33	50.8	311	2	US-10-087-013-10	Sequence 10, Appl							
433	33	50.8	311	2	US-10-153-273-21	Sequence 21, Appl							
434	33	50.8	312	2	US-08-894-017-10	Sequence 10, Appl							
435	33	50.8	312	2	US-09-456-474-10	Sequence 10, Appl							
436	33	50.8	312	2	US-09-328-352-6404	Sequence 6404, Ap							
437	33	50.8	323	2	US-09-252-991A-30565	Sequence 30565, A							
438	33	50.8	324	1	US-08-568-459A-17	Sequence 17, Appl							
439	33	50.8	324	1	US-08-487-826B-29	Sequence 29, Appl							
440	33	50.8	324	2	US-09-210-288-17	Sequence 17, Appl							
441	33	50.8	324	2	US-10-153-273-17	Sequence 17, Appl							
442	33	50.8	333	2	US-09-171-710-2	Sequence 2, Appl							
443	33	50.8	337	2	US-09-328-352-5787	Sequence 5787, Ap							
444	33	50.8	341	2	US-09-902-540-13921	Sequence 13921, A							
445	33	50.8	340	2	US-09-252-991A-27871	Sequence 27871, A							
446	33	50.8	350	2	US-09-270-767-46700	Sequence 46700, A							
447	33	50.8	352	2	US-09-252-991A-29664	Sequence 29664, A							
448	33	50.8	364	2	US-09-252-991A-19037	Sequence 19037, A							
449	33	50.8	391	2	US-09-198-452A-364	Sequence 364, App							
450	33	50.8	402	6	5352575-5	Patent No. 5352575							
451	33	50.8	404	2	US-09-506-286B-14	Sequence 14, Appl							
452	33	50.8	404	2	US-09-506-286B-17	Sequence 17, Appl							
453	33	50.8	404	2	US-09-762-861B-17	Sequence 17, Appl							
454	33	50.8	404	2	US-09-762-861B-17	Sequence 17, Appl							
455	33	50.8	404	2	US-10-065-133A-14	Sequence 14, Appl							
456	33	50.8	404	2	US-10-065-133A-17	Sequence 17, Appl							
457	33	50.8	404	2	US-10-434-811A-14	Sequence 14, Appl							
458	33	50.8	404	2	US-10-434-811A-17	Sequence 17, Appl							
459	33	50.8	407	2	US-09-248-796A-26731	Sequence 26731, A							
460	33	50.8	410	1	US-08-190-802A-46	Sequence 46, Appl							
461	33	50.8	410	2	US-08-477-346-46	Sequence 46, Appl							
462	33	50.8	410	2	US-08-473-089-46	Sequence 46, Appl							
463	33	50.8	410	2	US-08-487-072A-46	Sequence 46, Appl							
464	33	50.8	412	2	US-09-252-991A-17613	Sequence 17613, A							
465	33	50.8	414	1	US-08-984-171-1	Sequence 1, Appl							

ALIGNMENTS

RESULT 1
US-09-949-016-8330
; Sequence 8330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8330
; LENGTH: 547
; TYPE: PR1
; ORGANISM: Human
; US-09-949-016-8330

Query Match 100.0%; Score 65; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YALWWSNGK 10
| | | | | | | | | |
Db 217 YALWWSNGK 226

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,280A
FILING DATE: 18-MARCH-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5767242man D.
REGISTRATION NUMBER: 30,946
REFERENCES/DOCKET NUMBER: LUD 5330.1
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-619-280A-2

Query Match 100.0%; Score 65; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
Db 210 YALWSPNGK 219
|||||

RESULT 4
US-08-940-391-2
Sequence 2, Application US/08940391
Patent No. 5965373
GENERAL INFORMATION:
APPLICANT: Zimmermann, Rainer; Park, John E.;
APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,391
FILING DATE: 01-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/619,280
FILING DATE: 18-MARCH-1996
APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5965373man D.
REGISTRATION NUMBER: 30,946
REFERENCES/DOCKET NUMBER: LUD 5330.1
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear

```


US-08-940-391-2

Query Match 100.0%; Score 65; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
Db 210 YALWSPNGK 219

RESULT 5

US-09-265-606-2

; Sequence 2, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-265-606-2

Query Match 100.0%; Score 65; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
Db 210 YALWSPNGK 219

RESULT 6

US-09-518-550-26

; Sequence 26, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
; APPLICANT: TRAVIS, James
; APPLICANT: FOTEMPA, Jan
; APPLICANT: BANBULA, Agnieszka

us-10-774-242a-4.rai

Query Match 92.3%; Score 60; DB 2; Length 761;
Best Local Similarity 90.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
Db 210 YALWSPNGK 219

RESULT 7

PCT-US93-07923-11

; Sequence 11, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

PCT-US93-07923-11

Query Match 81.5%; Score 53; DB 4; Length 593;
Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPPG 9

Db 213 ALWWSPPG 220

RESULT 8

PCT-US93-07923-3
; Sequence 3, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlosman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

PCT-US93-07923-3

Query Match 81.5%; Score 53; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPPG 9

Db 202 ALWWSPPG 209

RESULT 9

PCT-US93-07923-2
; Sequence 2, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlosman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

PCT-US93-07923-2

Query Match 81.5%; Score 53; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWSPPG 9

Db 206 ALWWSPPG 213

RESULT 10

US-08-230-491A-3
; Sequence 3, Application US/08230491A
; Patent No. 5587299
; GENERAL INFORMATION:
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FELFE & LYNCH
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10022

COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT - ASC II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,491A
; FILING DATE: 20-APRIL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587299man D.

```

; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-230-491A-3

Query Match 81.5%; Score 53; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWMSPNG 9
Db 213 ALWMSPNG 220

RESULT 11
US-08-619-280A-3
; Sequence 3, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-619-280A-3

Query Match 81.5%; Score 53; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWMSPNG 9
Db 213 ALWMSPNG 220

RESULT 12
US-08-940-391-3
; Sequence 3, Application US/08940391
; Patent No. 595373
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
; TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-940-391-3

Query Match 81.5%; Score 53; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWMSPNG 9
Db 213 ALWMSPNG 220

RESULT 13
US-09-794-236-1
; Sequence 1, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-1
```

Query Match 81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 213 ALWSPNG 220
|||||

RESULT 14
US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. 6586198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 213 ALWSPNG 220
|||||

RESULT 15
US-09-949-016-6146
; Sequence 6146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6146
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6146

Query Match 81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 213 ALWSPNG 220
|||||

RESULT 16
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-1884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

Query Match 81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 213 ALWSPNG 220
|||||

RESULT 17
US-09-518-550-27
; Sequence 27, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
; APPLICANT: POTEMPA, Jan
; APPLICANT: TRAVIS, James
; APPLICANT: BANEULA, Agnieszka
; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
; FILE REFERENCE: 235.00190101
; CURRENT APPLICATION NUMBER: US/09/518,550
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,148
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US00/05551
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 48

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-550-27

Query Match      81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
DB      213 ALWWSPPNG 220

RESULT 18
US-10-423-714-6
; Sequence 6, Application US/10423714
; Patent No. 6887679
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
; FILE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. 6887679 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/423,714
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-423-714-6

Query Match      81.5%; Score 53; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
DB      213 ALWWSPPNG 220

RESULT 19
US-09-949-016-10450
; Sequence 10450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10450
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10450

Query Match      81.5%; Score 53; DB 2; Length 775;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
DB      213 ALWWSPPNG 220

RESULT 20
US-09-462-284-2
; Sequence 2, Application US/09462284
; Patent No. 6309868
; GENERAL INFORMATION:
; APPLICANT: Nestec S.A.
; APPLICANT: Monod, Michel
; APPLICANT: Doumas, Agnes
; APPLICANT: Affolter, Micheal
; APPLICANT: Van Den Broek, Peter
; TITLE OF INVENTION: CLONING OF THE
; FILE OF INVENTION: PROLYL-DIPEPTIDYL-PEPTIDASE FROM
; FILE REFERENCE: ASPERGILLUS ORYZAE
; CURRENT APPLICATION NUMBER: US/09/462,284
; CURRENT FILING DATE: 2000-01-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Fungus
US-09-462-284-2

Query Match      70.8%; Score 46; DB 2; Length 771;
Best Local Similarity 70.0%; Pred. No. 24;
Matches      7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 YALWWSPPNGK 10
DB      205 YALWFSPPDGE 214

RESULT 21
US-09-079-592-2
; Sequence 2, Application US/09079592B
; Patent No. 6664092
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Kimberly Brown
; APPLICANT: Michael W. Key
; APPLICANT: Alan Klotz
; APPLICANT: Tony Byun
; TITLE OF INVENTION: Polypeptides Having Dipeptidyl
; FILE OF INVENTION: Aminopeptidase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: S254.200-US
; CURRENT APPLICATION NUMBER: US/09/079,592B
; CURRENT FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 08/857,884
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: 60/062,892
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Aspergillus oryzae
US-09-079-592-2

Query Match      70.8%; Score 46; DB 2; Length 771;
Best Local Similarity 70.0%; Pred. No. 24;
Matches      7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 YALWWSPPNGK 10
DB      205 YALWFSPPDGE 214

```

```

; MOLECULE TYPE: peptide
US-08-838-543-4

Query Match      67.7%; Score 44; DB 1; Length 323;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 YALWSPNGK 10
DB      : ||| |||
        90 FRLWIPK 99

RESULT 24
US-08-704-711A-1
; Sequence 1, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; APPLICANT: HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-704-711A-1

Query Match      66.2%; Score 43; DB 2; Length 579;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWSPNGK 10
DB      : ||| |||
        415 ALFWPNGK 423

RESULT 25
```

```

; MOLECULE TYPE: peptide
US-09-902-540-11805

Query Match      69.2%; Score 45; DB 2; Length 574;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 WMSPNGK 10
DB      : ||| |||
        61 WMSPDGK 67

RESULT 23
US-08-838-543-4
; Sequence 4, Application US/08838543
; Patent No. 5954623
; GENERAL INFORMATION:
; APPLICANT: KREBBERS, ENNO
; APPLICANT: BROGLIE, KAREN E.
; TITLE OF INVENTION: CORN 4-(-GLUCANOTRANSFERASE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DUPONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (VER. 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,543
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MAJARIAN, WILLIAM R.
; REGISTRATION NUMBER: P-41,173
; REFERENCE/DOCKET NUMBER: BB-1101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4926
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
US-09-521-220-1
; Sequence 1, Application US/09521220
; Patent No. 6399348
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; METALLOPROTEASES, THEIR PRODUCTION AND USE
;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/521,220
; FILING DATE: 08-Mar-2000
; CLASSIFICATION: <unknown>
; 21-OCT-1994
; 17-MAR-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,711
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-521-220-1
Query Match 66.2%; Score 43; DB 2; Length 579;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALWSPNGK 10
Db 415 ALFWMPNGK 423

RESULT 26
US-08-704-711A-2
; Sequence 2, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
```

```
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-704-711A-2
Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALWSPNGK 10
Db 418 ALFWMPNGK 426

RESULT 27
US-08-448-489-1
; Sequence 1, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-250P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-448-489-1
Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ALWSPNGK 10
Db 418 ALFWMPNGK 10
```


Db 418 ALFWMPNGK 426

RESULT 28
US-09-211-704A-9
; Sequence 9, Application US/09211704A
; Patent No. 6271014
; GENERAL INFORMATION:
; APPLICANT: de Saint-Vin, Blandine Marie
; APPLICANT: Fossiez, Francois
; APPLICANT: Caux, Christophe
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Proteinases; Related Reagents
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNA Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,704A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/005,263
; FILING DATE: 09-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SP0781K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-211-704A-9

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
||| ||||
Db 418 ALFWMPNGK 426

RESULT 29
US-09-521-220-2
; Sequence 2, Application US/09521220
; Patent No. 6399348
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; HINZMANN, Bernd
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/521,220
; FILING DATE: 08-Mar-2000
; CLASSIFICATION: <Unknown>
; 21-OCT-1994
; 17-MAR-1994

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,711
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-521-220-2

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
||| ||||
Db 418 ALFWMPNGK 426

RESULT 30
US-09-391-104-28
; Sequence 28, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE.
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-391-104-28

Query Match 66.2%; Score 43; DB 2; Length 582;

```
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 2 ALWSPNGK 10
   ||| |||
Db 418 ALFWMPNGK 426

RESULT 31
US-09-919-497-84
; Sequence 84, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-84

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 2 ALWSPNGK 10
   ||| |||
Db 418 ALFWMPNGK 426

RESULT 32
US-09-689-730-1
; Sequence 1, Application US/09689730
; Patent No. 6825024
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290F
; CURRENT APPLICATION NUMBER: US/09/689,730
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US/08/448,489
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-689-730-1

Query Match 66.2%; Score 43; DB 2; Length 582;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 2 ALWSPNGK 10
   ||| |||
Db 418 ALFWMPNGK 426

RESULT 33
US-09-248-796A-18040
; Sequence 18040, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
```

```
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18040
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18040

Query Match 66.2%; Score 43; DB 2; Length 612;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LWMSPN 8
   ||| |||
Db 302 LWMSPN 307

RESULT 34
US-09-198-452A-702
; Sequence 702, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 702
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-702

Query Match 64.6%; Score 42; DB 2; Length 352;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YALWWSPNG 9
   : ||| |||
Db 124 FALWWSGLNG 132

RESULT 35
US-09-248-796A-19923
; Sequence 1923, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19923
; LENGTH: 396
; TYPE: PRT
```

```

; ORGANISM: Candida albicans
US-09-248-796A-19923

Query Match          64.6%; Score 42; DB 2; Length 396;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      3 LWSPNG 9
      :||||:|
Db      330 LWSPSG 336

RESULT 36
US-09-543-681A-5821
; Sequence 5821, Application US/09543681A
; Patent No. 6805709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5821
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5821

Query Match          64.6%; Score 42; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 VALMSP 7
      :|||||
Db      48 YALMWLP 54

RESULT 37
US-09-107-532A-7119
; Sequence 7119, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489

; ORGANISM: Enterococcus faecium
US-09-107-532A-7119
; Sequence 7119, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489

; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-667
; Sequence 667, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kaiman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 667
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0665
US-09-438-185A-667

Query Match          64.6%; Score 42; DB 2; Length 456;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 YALMWSPNG 9
      :|||||
Db      125 FALMWGLNG 133

RESULT 39
US-09-976-674-43
; Sequence 43, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
```

```
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-43

Query Match          64.6%; Score 42; DB 2; Length 691;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ALWSPNGK 10
| |||||:|
Db      151 AHWSPDGE 159

RESULT 40
US-09-976-674-41
; Sequence 41, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-41

Query Match          64.6%; Score 42; DB 2; Length 706;
Best Local Similarity 66.7%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ALWSPNGK 10
| |||||:|
Db      151 AHWSPDGE 159

RESULT 41
US-10-363-937-4
; Sequence 4, Application US/10363937
; Patent No. 6927056
; GENERAL INFORMATION:
; APPLICANT: Todd, Stephen; Delegeane, Angelo M.;
; APPLICANT: Gandhi, Ameena R.; Nguyen, Danniell B.;
; APPLICANT: Hafalia, April J.A.; Kearney, Liam;
; APPLICANT: Lu, Yan; Lee, Ernestine A.;
; APPLICANT: Chawla, Narinder K.; Das, Debopriya;
; APPLICANT: Arizuz, Chandra S.; Yao, Monique G.;
; APPLICANT: Kallick, Deborah A.; Elliott, Vicki S.;
; APPLICANT: Ding, Li; Yue, Henry;
; APPLICANT: Reddy, Roopa; Burford, Neil;
; APPLICANT: Baughn, Mariah R.; Lal, Preeti G.;
; APPLICANT: Borowsky, Mark L.; Lu, Dyung Aina M.;
; APPLICANT: Ramkumar, Jayalaxmi; Yang, Junming;
```

```
; APPLICANT: Tribouley, Catherine M.; Khan, Farrah A.;
; APPLICANT: Gururajan, Rajagopal; Tang, Y. Tom;
; APPLICANT: Au-Young, Janice; Warren, Bridget A.;
; APPLICANT: Hernandez, Roberto; Duggan, Brendan M.
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0212 USN
; CURRENT APPLICATION NUMBER: US/10/363,937
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: PCT/US01/28161
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 60/231,039
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,812
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/234,850
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236,500
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,773
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 60/239,658
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6927056 2923874CD1
US-10-363-937-4

Query Match          64.6%; Score 42; DB 2; Length 743;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ALWSPNGK 10
| |||||:|
Db      188 AHWSPDGE 196

RESULT 42
US-09-976-674-5
; Sequence 5, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 796
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-5

Query Match          64.6%; Score 42; DB 2; Length 796;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ALWSPNGK 10
| |||||:|
Db      241 AHWSPDGE 249
```

```
RESULT 43
US-09-462-845-3
; Sequence 3, Application US/09462845
; Patent No. 6723550
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/09/462,845
; CURRENT FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-462-845-3

Query Match      64.6%; Score 42; DB 2; Length 818;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALWSPNG 9
      | | | | |
Db      250 AAWSPTG 257

RESULT 44
US-10-402-312-3
; Sequence 3, Application US/10402312
; Patent No. 6833261
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/402,312
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-402-312-3

Query Match      64.6%; Score 42; DB 2; Length 818;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALWSPNG 9
      | | | | |
Db      250 AAWSPTG 257

RESULT 45
US-10-401-437-3
; Sequence 3, Application US/10401437
; Patent No. 6849440
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/401,437
; CURRENT FILING DATE: 2003-03-26
```

```
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-401-437-3

Query Match      64.6%; Score 42; DB 2; Length 818;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALWSPNG 9
      | | | | |
Db      250 AAWSPTG 257

RESULT 46
US-10-402-067-3
; Sequence 3, Application US/10402067
; Patent No. 6881562
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/402,067
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-402-067-3

Query Match      64.6%; Score 42; DB 2; Length 818;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 ALWSPNG 9
      | | | | |
Db      250 AAWSPTG 257

RESULT 47
US-10-401-436-3
; Sequence 3, Application US/10401436
; Patent No. 6911333
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC382-US
; CURRENT APPLICATION NUMBER: US/10/401,436
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: PCT/US98/14647
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: EP 97305237.7
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-401-436-3
```

Query Match 64.6%; Score 42; DB 2; Length 818;
Best Local Similarity 75.0%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
| | | | |
Db 250 AAWSPGTG 257

RESULT 48
US-09-248-796A-19283
; Sequence 19283, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19283
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19283

Query Match 63.1%; Score 41; DB 2; Length 320;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YALWSPNG 9
| | | | |
Db 178 YCLWSPNG 186

RESULT 49
US-09-211-704A-8
; Sequence 8, Application US/09211704A
; Patent No. 6271014
; GENERAL INFORMATION:
; APPLICANT: de Saint-Via, Blandine Marie
; APPLICANT: Fossiez, Francois
; APPLICANT: Caux, Christophe
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Proteinases; Related Reagents
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,704A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/005,263
; FILING DATE: 09-JAN-1998
; ATTORNEY/AGENT INFORMATION:

Query Match 63.1%; Score 41; DB 2; Length 564;
Best Local Similarity 62.5%; Pred. No. 1.le+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
| | | | |
Db 364 ALWWEPTG 371

RESULT 50
US-08-704-711A-3
; Sequence 3, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; APPLICANT: HINZMANN, Bernd
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-704-711A-3

Query Match 63.1%; Score 41; DB 2; Length 669;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 469 AIWWEPTG 476

RESULT 51

US-09-521-220-3
; Sequence 3, Application US/09521220
; Patent No. 6399348
; GENERAL INFORMATION:
; APPLICANT: WILL, Horst
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/521,220
; APPLICATION NUMBER: US/09/521,220
; FILING DATE: 08-Mar-2000
; CLASSIFICATION: <Unknown>
; 21-OCT-1994
; 17-MAR-1994

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,711
; FILING DATE: <Unknown>
; APPLICATION NUMBER: DE 4438838.1
; FILING DATE: 21-OCT-1994
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136

SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-521-220-3

Query Match 63.1%; Score 41; DB 2; Length 669;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 469 AIWWEPTG 476

RESULT 52

US-09-391-104-29

; Sequence 29, Application US/09391104
; Patent No. 6399371
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,
; PROTEINS ENCODED THEREFROM AND METHODS
; TITLE OF INVENTION: OF USING SAME
; FILE REFERENCE: 6073 US P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-391-104-29

Query Match 63.1%; Score 41; DB 2; Length 669;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
Db 469 AIWWEPTG 476

RESULT 53

US-09-902-540-10963
; Sequence 10963, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10963
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10963

Query Match 61.5%; Score 40; DB 2; Length 178;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LWSPNGK 10
Db 90 VWWSANGR 97

RESULT 54

US-09-449-218D-12
; Sequence 12, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.


```
; APPLICANT: Paeper, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 61.5%; Score 40; DB 2; Length 211;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WSPNG 9
Db 122 WRPNG 127

RESULT 55
US-09-668-529A-12
; Sequence 12, Application US/09668529A
; Patent No. 6489445
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paeper, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE
; FILE REFERENCE: 240083.508D1
; CURRENT APPLICATION NUMBER: US/09/668,529A
; CURRENT FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-668-529A-12

Query Match 61.5%; Score 40; DB 2; Length 211;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WSPNG 9
Db 122 WRPNG 127

RESULT 56
US-09-668-037A-12
; Sequence 12, Application US/09668037A
; Patent No. 6495736
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paeper, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-668-037A-12

Query Match 61.5%; Score 40; DB 2; Length 211;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WSPNG 9
Db 122 WRPNG 127

RESULT 57
US-09-668-021-12
; Sequence 12, Application US/09668021
; Patent No. 6803453
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paeper, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/668,021
; CURRENT FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 09/449,218
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-668-021-12

Query Match 61.5%; Score 40; DB 2; Length 211;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WSPNG 9
Db 122 WRPNG 127

RESULT 58
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paeper, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
```

; ORGANISM: Rattus norvegicus
US-09-668-037A-14
Query Match 61.5%; Score 40; DB 2; Length 213;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WSPNG 9
Db 124 WRPNG 129
Search completed: March 11, 2006, 12:27:17
Job time : 27.3684 secs

; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14
Query Match 61.5%; Score 40; DB 2; Length 213;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WSPNG 9
Db 124 WRPNG 129

RESULT 59
US-09-668-529A-14
; Sequence 14, Application US/09668529A
; Patent No. 6489445
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE
; FILE REFERENCE: 240083.508D1
; CURRENT APPLICATION NUMBER: US/09/668,529A
; CURRENT FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-668-529A-14

Query Match 61.5%; Score 40; DB 2; Length 213;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WSPNG 9
Db 124 WRPNG 129
RESULT 60
US-09-668-037A-14
; Sequence 14, Application US/09668037A
; Patent No. 6495736
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING BONE
; FILE REFERENCE: 240083.508D4
; CURRENT APPLICATION NUMBER: US/09/668,037A
; CURRENT FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT

Query Match 61.5%; Score 40; DB 2; Length 213;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WSPNG 9
Db 124 WRPNG 129

Query Match 61.5%; Score 40; DB 2; Length 213;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 WSPNG 9
Db 124 WRPNG 129

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:26:08 ; Search time 78.4211 Seconds
(without alignments)
53.280 Million cell updates/sec

Title: US-10-774-242A-4
Perfect score: 65
Sequence: 1 YALWSPNGK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	65	100.0	10	4	US-10-774-242-4
2	65	100.0	51	3	US-09-864-761-46651
3	65	100.0	57	4	US-10-029-386-29292
4	65	100.0	760	3	US-09-265-606-2
5	65	100.0	760	4	US-10-177-293-136
6	65	100.0	760	4	US-10-301-822-55
7	65	100.0	760	5	US-10-723-860-4171
8	65	100.0	760	5	US-10-884-070A-13
9	55	84.6	504	4	US-10-072-012-863
10	53	81.5	48	5	US-10-476-264-127
11	53	81.5	766	3	US-09-265-606-3
12	53	81.5	766	3	US-09-993-959-1
13	53	81.5	766	4	US-10-002-593-6
14	53	81.5	766	4	US-10-165-603-7
15	53	81.5	766	4	US-10-423-714-6
16	53	81.5	766	4	US-10-295-027-590
17	53	81.5	766	4	US-10-295-027-922
18	53	81.5	766	4	US-10-794-899-41
19	53	81.5	766	5	US-10-952-459-18
20	53	81.5	766	5	US-10-476-264-147
21	53	81.5	766	5	US-10-770-712-1
22	53	81.5	766	5	US-10-631-467-565
23	53	81.5	766	6	US-11-041-674-6
24	53	81.5	767	4	US-10-165-603-4
25	53	81.5	767	4	US-10-794-899-38
26	53	81.5	767	5	US-10-770-712-2
27	49	75.4	404	4	US-10-156-761-13020

28	47	72.3	760	5	US-10-770-712-3	Sequence 3, Appli
29	47	72.3	760	5	US-10-631-467-1390	Sequence 1390, Ap
30	47	72.3	829	4	US-10-369-493-7006	Sequence 7006, Ap
31	46	70.8	771	4	US-10-723-807-2	Sequence 2, Appli
32	45	69.2	661	4	US-10-369-493-19475	Sequence 19475, A
33	45	69.2	758	4	US-10-369-493-2301	Sequence 2301, Ap
34	44	67.7	323	4	US-10-336-597-4	Sequence 4, Appli
35	44	67.7	1256	4	US-10-156-761-9740	Sequence 9740, Ap
36	43.5	66.9	87	4	US-10-425-115-305249	Sequence 305249, A
37	43	66.2	112	5	US-10-450-763-40309	Sequence 40309, A
38	43	66.2	189	4	US-10-432-198-4	Sequence 4, Appli
39	43	66.2	189	4	US-10-432-198-5	Sequence 5, Appli
40	43	66.2	189	4	US-10-432-198-6	Sequence 6, Appli
41	43	66.2	189	4	US-10-432-198-7	Sequence 7, Appli
42	43	66.2	218	4	US-10-432-198-7	Sequence 7, Appli
43	43	66.2	280	4	US-10-424-599-225890	Sequence 225890, A
44	43	66.2	381	4	US-10-276-272A-10	Sequence 10, Appl
45	43	66.2	389	4	US-10-276-272A-4	Sequence 4, Appli
46	43	66.2	425	4	US-10-276-272A-6	Sequence 6, Appli
47	43	66.2	471	6	US-11-021-951-185	Sequence 185, App
48	43	66.2	582	3	US-09-916-658-4	Sequence 4, Appli
49	43	66.2	582	3	US-09-391-104-28	Sequence 28, Appl
50	43	66.2	582	3	US-09-801-196-27	Sequence 27, Appl
51	43	66.2	582	3	US-09-919-497-84	Sequence 84, Appl
52	43	66.2	582	3	US-09-916-849A-2	Sequence 2, Appli
53	43	66.2	582	4	US-10-133-797-6	Sequence 6, Appli
54	43	66.2	582	4	US-10-131-985-43	Sequence 43, Appl
55	43	66.2	582	4	US-10-411-010-26	Sequence 26, Appl
56	43	66.2	582	4	US-10-447-315-9	Sequence 9, Appli
57	43	66.2	582	4	US-10-276-272A-19	Sequence 19, Appl
58	43	66.2	582	5	US-10-901-417-43	Sequence 43, Appl
59	43	66.2	582	5	US-10-953-264-26	Sequence 26, Appl
60	43	66.2	582	5	US-10-773-446-100	Sequence 100, App
61	43	66.2	590	4	US-10-276-272A-2	Sequence 2, Appli
62	43	66.2	592	4	US-10-276-272A-12	Sequence 12, Appl
63	43	66.2	729	4	US-10-276-272A-15	Sequence 15, Appl
64	43	66.2	930	4	US-10-276-272A-13	Sequence 13, Appl
65	42	64.6	48	5	US-10-476-264-126	Sequence 126, App
66	42	64.6	64	4	US-10-437-963-122354	Sequence 122354, A
67	42	64.6	69	4	US-10-425-115-332920	Sequence 332920, A
68	42	64.6	352	4	US-10-289-762-702	Sequence 702, App
69	42	64.6	380	5	US-10-476-264-32	Sequence 32, Appl
70	42	64.6	438	4	US-10-282-122A-57752	Sequence 57752, A
71	42	64.6	579	4	US-10-369-493-3190	Sequence 3190, Ap
72	42	64.6	664	4	US-10-369-493-8177	Sequence 8177, Ap
73	42	64.6	691	3	US-09-976-674-43	Sequence 43, Appl
74	42	64.6	691	5	US-10-982-512-43	Sequence 43, Appl
75	42	64.6	706	3	US-09-976-674-41	Sequence 41, Appl
76	42	64.6	706	5	US-10-982-512-41	Sequence 41, Appl
77	42	64.6	743	4	US-10-363-937-4	Sequence 4, Appli
78	42	64.6	743	6	US-11-167-040-4	Sequence 4, Appli
79	42	64.6	746	5	US-10-476-264-69	Sequence 69, Appl
80	42	64.6	746	5	US-10-476-264-70	Sequence 70, Appl
81	42	64.6	746	5	US-10-476-264-72	Sequence 72, Appl
82	42	64.6	775	5	US-10-926-180A-35	Sequence 35, Appl
83	42	64.6	776	5	US-10-476-264-34	Sequence 34, Appl
84	42	64.6	789	5	US-10-476-264-48	Sequence 48, Appl
85	42	64.6	789	5	US-10-476-264-71	Sequence 71, Appl
86	42	64.6	796	3	US-09-976-674-5	Sequence 5, Appli
87	42	64.6	796	3	US-09-870-133-2	Sequence 2, Appli
88	42	64.6	796	4	US-10-160-501-5	Sequence 5, Appli
89	42	64.6	796	5	US-10-982-512-5	Sequence 5, Appli
90	42	64.6	796	5	US-10-433-757-2	Sequence 2, Appli
91	42	64.6	796	5	US-10-476-264-41	Sequence 41, Appl
92	42	64.6	796	5	US-10-476-264-47	Sequence 47, Appl
93	42	64.6	796	5	US-10-476-264-68	Sequence 68, Appl
94	42	64.6	796	5	US-10-476-264-106	Sequence 106, App
95	42	64.6	796	5	US-10-476-264-110	Sequence 110, App
96	42	64.6	796	5	US-10-476-264-142	Sequence 142, App
97	42	64.6	796	5	US-10-476-264-143	Sequence 143, App
98	42	64.6	796	5	US-10-476-264-146	Sequence 146, App
99	42	64.6	796	5	US-10-476-264-421	Sequence 421, App
100	42	64.6	797	5	US-10-476-264-36	Sequence 36, Appl

101	42	64.6	797	5	US-10-476-264-38	Sequence 38, Appl	174	40	61.5	258	5	US-10-732-923-16757	Sequence 16757, A
102	42	64.6	798	4	US-10-210-130-120	Sequence 120, Appl	175	40	61.5	285	5	US-10-739-930-7518	Sequence 7518, Ap
103	42	64.6	800	5	US-10-476-264-49	Sequence 49, Appl	176	40	61.5	288	3	US-09-883-249-2	Sequence 2, Appli
104	42	64.6	800	5	US-10-476-264-50	Sequence 50, Appl	177	40	61.5	289	3	US-10-732-923-6598	Sequence 6598, Ap
105	42	64.6	818	4	US-10-401-437-3	Sequence 3, Appli	178	40	61.5	329	4	US-10-208-018-104	Sequence 104, App
106	42	64.6	818	4	US-10-402-312-3	Sequence 3, Appli	179	40	61.5	335	4	US-10-282-122A-73122	Sequence 73122, A
107	42	64.6	818	4	US-10-402-067-3	Sequence 3, Appli	180	40	61.5	341	4	US-10-424-599-232753	Sequence 232753,
108	42	64.6	818	4	US-10-401-436-3	Sequence 3, Appli	181	40	61.5	406	4	US-10-374-780A-1805	Sequence 1805, Ap
109	42	64.6	818	4	US-10-369-493-22029	Sequence 22029, A	182	40	61.5	406	4	US-10-437-963-118768	Sequence 118768,
110	42	64.6	818	6	US-11-014-051-3	Sequence 3, Appli	183	40	61.5	408	5	US-10-732-923-65752	Sequence 65752, Ap
111	42	64.6	818	6	US-11-013-991-3	Sequence 3, Appli	184	40	61.5	416	4	US-10-424-599-274532	Sequence 274532,
112	42	64.6	818	6	US-11-014-339-3	Sequence 3, Appli	185	40	61.5	421	4	US-10-425-115-251099	Sequence 251099,
113	42	64.6	818	6	US-11-014-364-3	Sequence 3, Appli	186	40	61.5	421	4	US-10-067-443-4	Sequence 4, Appli
114	42	64.6	818	6	US-11-014-432-3	Sequence 3, Appli	187	40	61.5	421	4	US-10-067-443-28	Sequence 4, Appli
115	41	63.1	63	4	US-10-029-386-29546	Sequence 29546, A	188	40	61.5	421	4	US-10-649-273-4	Sequence 28, Appl
116	41	63.1	142	4	US-10-425-115-238885	Sequence 238885, A	189	40	61.5	421	4	US-10-649-273-28	Sequence 4, Appli
117	41	63.1	193	4	US-10-425-115-194175	Sequence 194175, A	190	40	61.5	421	4	US-10-651-722-4	Sequence 28, Appl
118	41	63.1	458	4	US-10-032-585-7575	Sequence 7575, Ap	191	40	61.5	421	4	US-10-651-722-28	Sequence 28, Appl
119	41	63.1	530	4	US-10-425-114-39737	Sequence 39737, A	192	40	61.5	431	4	US-10-437-963-125484	Sequence 51343, A
120	41	63.1	557	5	US-10-659-004-44	Sequence 44, Appl	193	40	61.5	471	4	US-10-282-122A-51343	Sequence 51343, A
121	41	63.1	669	3	US-09-391-104-29	Sequence 29, Appl	194	40	61.5	538	4	US-10-282-122A-73198	Sequence 73198, A
122	41	63.1	669	3	US-09-801-196-28	Sequence 28, Appl	195	40	61.5	539	3	US-09-815-242-13809	Sequence 13809, A
123	41	63.1	669	4	US-10-131-985-45	Sequence 45, Appl	196	40	61.5	539	4	US-10-282-122A-75616	Sequence 75616, A
124	41	63.1	669	5	US-10-901-417-45	Sequence 45, Appl	197	40	61.5	816	4	US-10-767-701-45650	Sequence 45650, A
125	41	63.1	803	4	US-10-156-761-12925	Sequence 12925, A	198	40	61.5	855	4	US-10-425-115-330355	Sequence 330355,
126	41	63.1	803	4	US-10-205-219-58	Sequence 59, Appl	199	40	61.5	1086	4	US-10-425-115-62437	Sequence 62437, A
127	41	63.1	805	5	US-10-476-264-145	Sequence 145, App	200	40	61.5	2314	4	US-10-437-963-185483	Sequence 185483,
128	41	63.1	865	5	US-10-476-264-144	Sequence 144, App	201	40	61.5	2329	5	US-10-469-865-5	Sequence 5, Appli
129	41	63.1	972	4	US-10-424-599-273839	Sequence 273839, A	202	40	61.5	2335	5	US-10-469-865-2	Sequence 2, Appli
130	41	63.1	1132	4	US-10-094-749-1920	Sequence 1920, Ap	203	40	61.5	2350	4	US-10-437-963-185481	Sequence 7, Appli
131	41	63.1	3065	5	US-10-486-977-19	Sequence 19, Appl	204	40	61.5	2352	5	US-10-469-865-7	Sequence 3504, Ap
132	40	61.5	18	4	US-10-463-190-76	Sequence 76, Appl	205	40	61.5	2356	4	US-10-128-714-3504	Sequence 6, Appli
133	40	61.5	18	5	US-10-868-497-31	Sequence 31, Appl	206	40	61.5	2391	5	US-10-469-865-6	Sequence 4, Appli
134	40	61.5	19	5	US-10-463-190-77	Sequence 77, Appl	207	40	61.5	2396	5	US-11-097-143-18933	Sequence 18933, A
135	40	61.5	19	5	US-10-868-497-32	Sequence 32, Appl	208	40	61.5	2396	6	US-10-032-585-7352	Sequence 8504, Ap
136	40	61.5	24	4	US-10-464-368-123	Sequence 123, App	209	40	61.5	2416	4	US-10-128-714-8504	Sequence 21, Appl
137	40	61.5	26	4	US-10-463-190-99	Sequence 99, Appl	210	39	60.0	3170	4	US-09-731-449-21	Sequence 21, Appl
138	40	61.5	26	5	US-10-868-497-54	Sequence 54, App	211	39	60.0	42	3	US-10-434-426-21	Sequence 21, Appl
139	40	61.5	47	3	US-09-782-980-57	Sequence 57, Appl	212	39	60.0	42	4	US-10-427-224-35	Sequence 23, Appl
140	40	61.5	47	4	US-10-806-018-57	Sequence 57, Appl	213	39	60.0	47	3	US-09-731-449-22	Sequence 23, Appl
141	40	61.5	52	4	US-10-464-368-120	Sequence 120, App	214	39	60.0	47	3	US-09-731-449-23	Sequence 17, Appl
142	40	61.5	66	4	US-10-424-599-216320	Sequence 216320, App	215	39	60.0	47	4	US-10-105-934-17	Sequence 22, Appl
143	40	61.5	135	5	US-10-779-597-106	Sequence 106, App	216	39	60.0	47	4	US-10-254-426-22	Sequence 23, Appl
144	40	61.5	185	3	US-09-867-274-4	Sequence 4, Appli	217	39	60.0	47	4	US-10-854-426-23	Sequence 17, Appl
145	40	61.5	185	4	US-10-679-670A-4	Sequence 4, Appli	218	39	60.0	47	5	US-10-895-676-17	Sequence 33734, A
146	40	61.5	188	5	US-10-463-190-143	Sequence 143, App	219	39	60.0	81	4	US-10-029-386-33734	Sequence 8248, Ap
147	40	61.5	188	5	US-10-868-497-112	Sequence 112, App	220	39	60.0	82	3	US-09-864-408A-8248	Sequence 54967, A
148	40	61.5	190	4	US-10-463-190-65	Sequence 65, Appl	221	39	60.0	93	4	US-10-767-701-54846	Sequence 254967,
149	40	61.5	190	5	US-10-868-497-20	Sequence 20, Appl	222	39	60.0	94	4	US-10-424-599-254967	Sequence 260424,
150	40	61.5	200	4	US-10-424-599-179787	Sequence 179787, A	223	39	60.0	97	4	US-10-424-599-260424	Sequence 8192, Ap
151	40	61.5	200	4	US-10-464-368-49	Sequence 49, Appl	224	39	60.0	97	4	US-10-732-923-8192	Sequence 330315,
152	40	61.5	208	3	US-09-867-274-6	Sequence 6, Appli	225	39	60.0	116	5	US-10-425-115-350315	Sequence 253195,
153	40	61.5	208	4	US-10-679-670A-6	Sequence 6, Appli	226	39	60.0	117	4	US-10-425-115-350195	Sequence 243926,
154	40	61.5	211	4	US-10-384-893-12	Sequence 12, Appl	227	39	60.0	119	4	US-10-424-599-243926	Sequence 302827,
155	40	61.5	211	4	US-10-377-315-6	Sequence 6, Appli	228	39	60.0	140	4	US-10-425-115-302827	Sequence 38007, A
156	40	61.5	211	4	US-10-463-190-12	Sequence 12, Appl	229	39	60.0	142	4	US-10-767-701-39007	Sequence 348988,
157	40	61.5	211	4	US-10-464-368-46	Sequence 46, Appl	230	39	60.0	164	4	US-10-425-115-348988	Sequence 38756, A
158	40	61.5	211	4	US-10-464-368-47	Sequence 47, Appl	231	39	60.0	191	4	US-10-767-701-38756	Sequence 199467,
159	40	61.5	211	4	US-10-464-368-48	Sequence 48, Appl	232	39	60.0	202	4	US-10-425-115-199471	Sequence 199471,
160	40	61.5	211	4	US-10-095-248A-12	Sequence 12, Appl	233	39	60.0	202	4	US-10-425-115-258199	Sequence 258199,
161	40	61.5	211	5	US-10-788-606-12	Sequence 12, Appl	234	39	60.0	204	4	US-10-425-115-67255	Sequence 348991,
162	40	61.5	211	5	US-10-799-162-12	Sequence 12, Appl	235	39	60.0	243	4	US-10-425-114-67255	Sequence 78255, A
163	40	61.5	211	5	US-10-868-497-66	Sequence 66, Appl	236	39	60.0	246	4	US-10-425-114-67255	Sequence 30386, A
164	40	61.5	213	4	US-10-384-893-14	Sequence 14, Appl	237	39	60.0	250	5	US-10-450-763-30386	Sequence 62233, A
165	40	61.5	213	4	US-10-377-315-8	Sequence 8, Appli	238	39	60.0	254	4	US-10-425-114-62233	Sequence 32397, A
166	40	61.5	213	4	US-10-463-190-14	Sequence 14, Appl	239	39	60.0	269	3	US-10-767-701-32397	Sequence 12, Appl
167	40	61.5	213	4	US-10-464-368-53	Sequence 53, Appl	240	39	60.0	305	3	US-09-988-067B-12	Sequence 323521,
168	40	61.5	213	4	US-10-095-248A-14	Sequence 14, Appl	241	39	60.0	314	4	US-10-425-115-323521	Sequence 5, Appli
169	40	61.5	213	4	US-10-788-606-14	Sequence 14, Appl	242	39	60.0	321	4	US-10-427-224-5	Sequence 256126,
170	40	61.5	213	5	US-10-799-162-14	Sequence 14, Appl	243	39	60.0	324	4	US-10-425-115-256126	Sequence 24, Appl
171	40	61.5	221	5	US-10-868-497-68	Sequence 68, Appl	244	39	60.0	332	4	US-10-631-581-24	Sequence 18, Appl
172	40	61.5	221	4	US-10-437-963-154853	Sequence 154853, A	245	39	60.0	333	4	US-10-631-581-18	
173	40	61.5	232	4	US-10-425-115-329758	Sequence 329758, A	246	39	60.0	333	4		

247	39	60.0	333	4	US-10-631-581-20	Sequence 20, Appl	320	38	58.5	485	4	US-10-132-744A-6	Sequence 6, Appl
248	39	60.0	336	4	US-10-437-963-136543	Sequence 136543, A	321	38	58.5	507	5	US-10-810-362-2	Sequence 2, Appl
249	39	60.0	338	4	US-10-425-114-59599	Sequence 59599, A	322	38	58.5	598	4	US-10-437-963-148984	Sequence 148984, A
250	39	60.0	370	4	US-10-437-963-166080	Sequence 166080, A	323	38	58.5	711	4	US-10-425-114-72929	Sequence 23019, A
251	39	60.0	371	4	US-10-425-114-39491	Sequence 39491, A	324	38	58.5	745	6	US-11-097-143-23019	Sequence 23019, A
252	39	60.0	399	4	US-10-369-493-2575	Sequence 2575, Ap	325	38	58.5	813	4	US-10-437-963-186828	Sequence 186828, A
253	39	60.0	465	4	US-10-369-493-22543	Sequence 22543, A	326	38	58.5	869	4	US-10-437-963-170106	Sequence 170106, A
254	39	60.0	525	4	US-10-424-599-268964	Sequence 268964, A	327	38	58.5	1025	4	US-10-437-963-170107	Sequence 170107, A
255	39	60.0	545	4	US-10-424-599-129624	Sequence 129624, A	328	38	58.5	1073	4	US-10-437-963-114026	Sequence 114026, A
256	39	60.0	575	4	US-10-425-115-188553	Sequence 188553, A	329	38	58.5	1163	4	US-10-437-963-151916	Sequence 151916, A
257	39	60.0	624	4	US-10-437-963-177572	Sequence 177572, A	330	37	56.9	44	4	US-10-767-701-58561	Sequence 58561, A
258	39	60.0	657	4	US-10-369-493-19349	Sequence 19349, A	331	37	56.9	48	4	US-10-425-115-19489	Sequence 19489, A
259	39	60.0	744	4	US-10-437-963-181621	Sequence 181621, A	332	37	56.9	50	4	US-10-424-599-272137	Sequence 272137, A
260	39	60.0	766	3	US-09-801-368-316	Sequence 316, App	333	37	56.9	52	4	US-10-424-599-272137	Sequence 272137, A
261	39	60.0	887	4	US-10-437-963-163583	Sequence 163583, A	334	37	56.9	63	4	US-10-425-115-265506	Sequence 265506, A
262	39	60.0	899	4	US-10-298-122-13	Sequence 13, Appl	335	37	56.9	66	4	US-10-437-963-153607	Sequence 153607, A
263	39	60.0	900	4	US-10-369-493-3756	Sequence 3756, Ap	336	37	56.9	67	4	US-10-424-599-173654	Sequence 173654, A
264	39	60.0	1018	4	US-10-437-963-122784	Sequence 122784, A	337	37	56.9	73	4	US-10-437-963-160390	Sequence 160390, A
265	39	60.0	1064	5	US-10-732-923-8191	Sequence 8191, Ap	338	37	56.9	77	3	US-10-437-963-170107	Sequence 170107, A
266	39	60.0	1064	5	US-10-732-923-8191	Sequence 8196, Ap	339	37	56.9	79	4	US-10-424-599-224541	Sequence 224541, A
267	39	60.0	1064	5	US-10-732-923-8294	Sequence 8294, Ap	340	37	56.9	79	5	US-10-450-763-57770	Sequence 57770, A
268	39	60.0	1065	5	US-10-732-923-8291	Sequence 8291, Ap	341	37	56.9	86	4	US-10-425-115-221398	Sequence 221398, A
269	39	60.0	1072	5	US-10-732-923-8209	Sequence 8209, Ap	342	37	56.9	86	4	US-10-425-115-221398	Sequence 221398, A
270	39	60.0	1072	5	US-10-732-923-8211	Sequence 8211, Ap	343	37	56.9	93	4	US-10-425-115-251505	Sequence 251505, A
271	39	60.0	1234	5	US-10-741-600-1357	Sequence 1357, Ap	344	37	56.9	107	4	US-10-425-115-251505	Sequence 251505, A
272	39	60.0	1365	5	US-10-741-600-1355	Sequence 1355, Ap	345	37	56.9	110	4	US-10-425-115-210087	Sequence 210087, A
273	39	60.0	1366	5	US-10-741-600-1355	Sequence 1355, Ap	346	37	56.9	120	4	US-10-425-115-210087	Sequence 210087, A
274	39	60.0	1411	5	US-10-741-600-1356	Sequence 1356, Ap	347	37	56.9	125	4	US-10-425-115-318397	Sequence 318397, A
275	39	60.0	2351	4	US-10-437-963-170608	Sequence 170608, A	348	37	56.9	128	4	US-10-181-611-22	Sequence 22, Appl
276	39	60.0	2505	4	US-10-437-963-170612	Sequence 170612, A	349	37	56.9	128	4	US-10-181-611-24	Sequence 24, Appl
277	38	58.5	8	4	US-10-367-580-243	Sequence 243, App	350	37	56.9	129	3	US-09-894-159-87	Sequence 87, Appl
278	38	58.5	8	4	US-10-367-593-243	Sequence 243, App	351	37	56.9	129	3	US-09-894-159-88	Sequence 88, Appl
279	38	58.5	8	4	US-10-367-594-243	Sequence 243, App	352	37	56.9	129	3	US-10-424-599-214396	Sequence 214396, A
280	38	58.5	8	4	US-10-367-654-243	Sequence 243, App	353	37	56.9	141	4	US-10-425-115-234382	Sequence 234382, A
281	38	58.5	8	4	US-10-367-654-243	Sequence 243, App	354	37	56.9	143	4	US-10-425-115-360675	Sequence 360675, A
282	38	58.5	8	4	US-10-367-658-243	Sequence 243, App	355	37	56.9	143	4	US-10-425-115-360675	Sequence 360675, A
283	38	58.5	8	4	US-10-367-674-243	Sequence 243, App	356	37	56.9	145	4	US-10-425-115-287052	Sequence 287052, A
284	38	58.5	8	5	US-10-820-067A-222	Sequence 168, App	357	37	56.9	147	3	US-09-894-159-85	Sequence 85, Appl
285	38	58.5	62	4	US-10-012-600B-168	Sequence 38016, A	358	37	56.9	149	4	US-10-425-114-41655	Sequence 41655, A
286	38	58.5	63	5	US-10-450-763-38016	Sequence 261570, A	359	37	56.9	149	4	US-10-425-114-41655	Sequence 41655, A
287	38	58.5	99	4	US-10-425-115-261570	Sequence 261570, A	360	37	56.9	155	4	US-10-425-114-72875	Sequence 72875, A
288	38	58.5	104	4	US-10-104-047-2562	Sequence 2562, Ap	361	37	56.9	155	4	US-10-437-963-135547	Sequence 135547, A
289	38	58.5	106	4	US-10-424-599-233520	Sequence 233520, A	362	37	56.9	162	4	US-10-437-963-191787	Sequence 191787, A
290	38	58.5	122	4	US-10-437-963-164359	Sequence 164359, A	363	37	56.9	168	4	US-10-467-595-7	Sequence 7, Appl
291	38	58.5	141	4	US-10-425-115-282084	Sequence 282084, A	364	37	56.9	191	4	US-10-437-963-146220	Sequence 146220, A
292	38	58.5	161	4	US-10-425-115-279646	Sequence 279646, A	365	37	56.9	219	4	US-10-425-115-354832	Sequence 354832, A
293	38	58.5	175	4	US-10-424-599-204681	Sequence 204681, A	366	37	56.9	227	4	US-10-437-963-190377	Sequence 190377, A
294	38	58.5	185	5	US-10-450-763-42653	Sequence 42653, A	367	37	56.9	237	4	US-10-425-114-65204	Sequence 65204, A
295	38	58.5	208	4	US-10-424-599-204680	Sequence 204680, A	368	37	56.9	240	4	US-10-437-963-144769	Sequence 144769, A
296	38	58.5	229	4	US-10-437-963-159784	Sequence 159784, A	369	37	56.9	257	4	US-10-425-115-264669	Sequence 264669, A
297	38	58.5	229	4	US-10-425-115-218475	Sequence 218475, A	370	37	56.9	274	5	US-10-739-930-9786	Sequence 9786, Ap
298	38	58.5	230	4	US-10-282-122A-69554	Sequence 69554, A	371	37	56.9	274	5	US-10-732-923-5831	Sequence 5831, Ap
299	38	58.5	232	4	US-10-282-122A-69554	Sequence 69554, A	372	37	56.9	281	4	US-10-437-963-190984	Sequence 190984, A
300	38	58.5	240	4	US-10-369-493-20236	Sequence 20236, A	373	37	56.9	285	4	US-10-437-963-203675	Sequence 203675, A
301	38	58.5	248	5	US-10-450-763-48261	Sequence 48261, A	374	37	56.9	285	4	US-10-223-074-5	Sequence 5, Appl
302	38	58.5	255	5	US-10-450-763-48087	Sequence 48087, A	375	37	56.9	292	5	US-10-450-763-46142	Sequence 46142, A
303	38	58.5	261	4	US-10-132-744A-4	Sequence 4, Appl	376	37	56.9	320	5	US-10-450-763-51704	Sequence 51704, A
304	38	58.5	306	4	US-10-424-599-264697	Sequence 264697, A	377	37	56.9	326	4	US-10-282-122A-55815	Sequence 55815, A
305	38	58.5	340	4	US-10-128-714-8037	Sequence 8037, Ap	378	37	56.9	329	4	US-10-425-115-216714	Sequence 216714, A
306	38	58.5	355	3	US-09-733-524-8	Sequence 8, Appl	379	37	56.9	332	4	US-10-437-963-126311	Sequence 126311, A
307	38	58.5	355	3	US-10-392-098-12	Sequence 12, Appl	380	37	56.9	332	4	US-10-425-114-53737	Sequence 53737, A
308	38	58.5	356	4	US-10-120-319-12	Sequence 12, Appl	381	37	56.9	333	4	US-10-425-114-53737	Sequence 53737, A
309	38	58.5	356	4	US-10-189-977-12	Sequence 12, Appl	382	37	56.9	336	5	US-10-450-763-47965	Sequence 47965, A
310	38	58.5	362	3	US-09-902-003-2	Sequence 2, Appl	383	37	56.9	396	4	US-10-425-114-40989	Sequence 40989, A
311	38	58.5	362	4	US-10-445-718-2	Sequence 2, Appl	384	37	56.9	397	4	US-10-437-963-184849	Sequence 184849, A
312	38	58.5	367	4	US-10-437-963-197841	Sequence 197841, A	385	37	56.9	406	3	US-09-731-872-284	Sequence 284, App
313	38	58.5	392	4	US-10-262-839-156	Sequence 156, App	386	37	56.9	406	3	US-09-876-997-284	Sequence 284, App
314	38	58.5	399	4	US-10-477-269-4	Sequence 4, Appl	387	37	56.9	406	3	US-10-643-836-284	Sequence 284, App
315	38	58.5	441	4	US-10-424-599-190978	Sequence 190978, A	388	37	56.9	406	5	US-10-767-701-47183	Sequence 47183, A
316	38	58.5	441	4	US-10-425-114-55268	Sequence 55268, A	389	37	56.9	415	4	US-10-425-115-217681	Sequence 217681, A
317	38	58.5	447	4	US-10-437-963-178074	Sequence 178074, A	390	37	56.9	415	4	US-10-425-115-342165	Sequence 342165, A
318	38	58.5	479	5	US-10-739-930-10964	Sequence 10964, A	391	37	56.9	419	4	US-10-425-114-63185	Sequence 63185, A
319	38	58.5	484	4	US-10-132-744A-2	Sequence 2, Appl	392	37	56.9	427	6	US-10-437-963-173383	Sequence 173383, A
										428	5	US-11-097-143-40869	Sequence 40869, A
										430	4	US-10-631-467-582	Sequence 582, App
												US-10-437-963-186065	Sequence 186065, A

393	37	56.9	452	4	US-10-425-114-59534	Sequence 69534, A	466	36	55.4	125	4	US-10-425-115-323310	Sequence 323310,
394	37	56.9	458	4	US-10-287-971-52	Sequence 52, Appl	467	36	55.4	128	4	US-10-425-115-284941	Sequence 284941,
395	37	56.9	475	4	US-10-437-963-191788	Sequence 191788, A	468	36	55.4	130	4	US-10-437-963-140418	Sequence 140418,
396	37	56.9	477	4	US-10-425-114-60716	Sequence 60716, A	469	36	55.4	131	3	US-09-908-322-60	Sequence 60, Appl
397	37	56.9	486	4	US-10-080-960-2	Sequence 2, Appl	470	36	55.4	131	3	US-09-783-931-60	Sequence 60, Appl
398	37	56.9	492	3	US-09-844-948-2	Sequence 2, Appl	471	36	55.4	139	4	US-10-001-857-186	Sequence 186, App
399	37	56.9	492	4	US-10-184-648-20	Sequence 20, Appl	472	36	55.4	140	4	US-10-424-599-154193	Sequence 154193,
400	37	56.9	508	3	US-09-016-159-5	Sequence 5, Appl	473	36	55.4	144	3	US-09-860-232A-20	Sequence 20, Appl
401	37	56.9	508	4	US-10-287-971-54	Sequence 54, Appl	474	36	55.4	147	4	US-10-424-599-225582	Sequence 225582,
402	37	56.9	508	5	US-10-684-206-24	Sequence 24, Appl	475	36	55.4	156	4	US-10-437-963-170096	Sequence 170096,
403	37	56.9	508	5	US-10-482-029-192	Sequence 192, App	476	36	55.4	170	4	US-10-425-115-257739	Sequence 257739,
404	37	56.9	530	4	US-10-437-963-125529	Sequence 125529, A	477	36	55.4	171	5	US-10-450-763-53099	Sequence 53099, A
405	37	56.9	539	4	US-10-282-122A-56178	Sequence 122A-56178, A	478	36	55.4	209	4	US-10-424-599-158741	Sequence 158741,
406	37	56.9	540	4	US-10-282-122A-60328	Sequence 60328, A	479	36	55.4	211	4	US-10-425-115-221741	Sequence 221741,
407	37	56.9	597	4	US-10-433-005-2	Sequence 2, Appl	480	36	55.4	213	3	US-09-989-442-90	Sequence 90, Appl
408	37	56.9	602	4	US-10-424-599-184579	Sequence 184579, A	481	36	55.4	215	4	US-10-425-114-49985	Sequence 49985, A
409	37	56.9	606	4	US-10-425-115-211037	Sequence 211037, A	482	36	55.4	226	4	US-10-369-493-9488	Sequence 9488, Ap
410	37	56.9	610	4	US-10-437-963-169590	Sequence 169590, A	483	36	55.4	231	4	US-10-282-122A-68142	Sequence 68142, A
411	37	56.9	610	4	US-10-437-963-169590	Sequence 169590, A	484	36	55.4	239	4	US-10-424-599-158739	Sequence 158739,
412	37	56.9	635	4	US-10-425-115-354836	Sequence 354836, A	485	36	55.4	257	3	US-09-860-232A-19	Sequence 19, Appl
413	37	56.9	687	4	US-10-425-114-61163	Sequence 61163, A	486	36	55.4	264	4	US-10-425-115-253657	Sequence 253657,
414	37	56.9	719	4	US-10-437-963-153848	Sequence 153848, A	487	36	55.4	278	4	US-10-437-963-143721	Sequence 143721,
415	37	56.9	749	4	US-10-424-599-174708	Sequence 174708, A	488	36	55.4	285	4	US-10-424-599-184021	Sequence 184021,
416	37	56.9	762	4	US-10-101-464A-114	Sequence 2376, A	489	36	55.4	285	4	US-10-424-599-22482	Sequence 22482, A
417	37	56.9	762	5	US-10-864-252-114	Sequence 114, App	490	36	55.4	303	6	US-11-097-143-22833	Sequence 22833, A
418	37	56.9	779	4	US-10-369-493-6623	Sequence 6623, Ap	491	36	55.4	337	6	US-11-097-143-20403	Sequence 20403, A
419	37	56.9	785	6	US-10-097-143-11397	Sequence 11397, A	492	36	55.4	365	4	US-10-369-493-22931	Sequence 22931, A
420	37	56.9	799	4	US-10-369-493-6622	Sequence 6622, Ap	493	36	55.4	385	4	US-10-369-493-19604	Sequence 19604, A
421	37	56.9	805	4	US-10-437-963-107258	Sequence 107258, A	494	36	55.4	398	4	US-10-282-122A-53157	Sequence 53157, A
422	37	56.9	897	6	US-11-097-143-25761	Sequence 25761, A	495	36	55.4	405	3	US-09-863-475A-8	Sequence 8, Appl
423	37	56.9	931	4	US-10-369-493-22397	Sequence 22397, A	496	36	55.4	405	4	US-10-234-041-11	Sequence 11, Appl
424	37	56.9	931	4	US-10-732-807-11	Sequence 11, Appl	497	36	55.4	410	4	US-10-156-761-11310	Sequence 11310, A
425	37	56.9	1163	5	US-10-732-923-8775	Sequence 8775, Ap	498	36	55.4	417	4	US-10-437-963-150752	Sequence 150752,
426	37	56.9	1166	5	US-10-732-923-8184	Sequence 8184, Ap	499	36	55.4	426	3	US-09-738-626-6036	Sequence 6036, Ap
427	37	56.9	1166	5	US-10-732-923-8184	Sequence 8184, Ap	500	36	55.4	426	4	US-10-627-476-424	Sequence 424, App
428	37	56.9	1166	5	US-10-732-923-8184	Sequence 8184, Ap							
429	37	56.9	1181	5	US-10-732-923-8300	Sequence 8300, Ap							
430	37	56.9	1534	5	US-10-966-482-28	Sequence 28, Appl							
431	36.5	56.2	1370	5	US-10-732-923-15002	Sequence 15002, A							
432	36.5	56.2	1370	5	US-10-732-923-15001	Sequence 15001, A							
433	36	55.4	30	5	US-10-450-763-47967	Sequence 47967, A							
434	36	55.4	47	5	US-10-450-763-31355	Sequence 31355, A							
435	36	55.4	57	4	US-10-029-386-30418	Sequence 30418, A							
436	36	55.4	57	4	US-10-425-115-357495	Sequence 357495, A							
437	36	55.4	68	4	US-10-424-599-190399	Sequence 190399, A							
438	36	55.4	70	3	US-09-864-761-40436	Sequence 40436, A							
439	36	55.4	71	4	US-10-425-115-213643	Sequence 213643, A							
440	36	55.4	73	4	US-10-425-115-283372	Sequence 283372, A							
441	36	55.4	78	3	US-09-925-297-752	Sequence 752, App							
442	36	55.4	82	4	US-10-425-115-192457	Sequence 192457, A							
443	36	55.4	82	4	US-10-424-599-195168	Sequence 195168, A							
444	36	55.4	83	3	US-09-908-711-103	Sequence 103, App							
445	36	55.4	83	4	US-09-764-891-3234	Sequence 3234, Ap							
446	36	55.4	83	4	US-10-437-963-172055	Sequence 172055, A							
447	36	55.4	84	4	US-10-424-599-160178	Sequence 160178, A							
448	36	55.4	85	4	US-10-097-065-183	Sequence 183, App							
449	36	55.4	85	4	US-10-372-876-183	Sequence 876, App							
450	36	55.4	86	4	US-10-316-194-48	Sequence 48, Appl							
451	36	55.4	86	4	US-10-425-115-242928	Sequence 242928, A							
452	36	55.4	86	5	US-10-732-180-48	Sequence 48, Appl							
453	36	55.4	92	4	US-10-425-115-254370	Sequence 254370, A							
454	36	55.4	95	4	US-10-424-599-864523	Sequence 864523, A							
455	36	55.4	99	4	US-10-425-115-302871	Sequence 302871, A							
456	36	55.4	99	5	US-10-732-180-233	Sequence 233, App							
457	36	55.4	100	4	US-10-425-115-198514	Sequence 198514, A							
458	36	55.4	101	4	US-10-316-194-46	Sequence 46, Appl							
459	36	55.4	101	5	US-10-732-180-46	Sequence 46, Appl							
460	36	55.4	103	4	US-10-425-115-242450	Sequence 242450, A							
461	36	55.4	103	4	US-10-425-115-307654	Sequence 307654, A							
462	36	55.4	103	5	US-10-732-180-261	Sequence 261, App							
463	36	55.4	106	4	US-10-425-114-70230	Sequence 70230, A							
464	36	55.4	110	4	US-10-424-599-246912	Sequence 246912, A							
465	36	55.4	116	5	US-10-450-763-56727	Sequence 56727, A							

ALIGNMENTS

RESULT 1

US-10-774-242-4
; Sequence 4, Application US/10774242
; Publication No. US20040203102A1
; GENERAL INFORMATION:
; APPLICANT: McKee, Patrick A
; APPLICANT: Lee, Kyung N.
; APPLICANT: Christiansen, Victoria J.
; TITLE OF INVENTION: ANTIPLASMIN CLEAVING ENZYME
; FILE REFERENCE: 5920.646
; CURRENT APPLICATION NUMBER: US/10/774,242
; PRIOR FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: 60/445,774
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-774-242-4

Query Match 100.0%; Score 65; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWWSNGK 10

Dd 1 YALWWSNGK 10

RESULT 2

US-09-864-761-46651
; Sequence 46651, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46651
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007750.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65
; OTHER INFORMATION: EST HUMAN HIT: AU136648.1, EVALUATE 1.00e-25
; OTHER INFORMATION: SWISSPROT HIT: P27487, EVALUATE 2.00e-16
US-09-864-761-46651

Query Match 100.0%; Score 65; DB 3; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YALWWSPNGK 10
|||||

Db 7 YALWWSPNGK 16

RESULT 3

US-10-029-386-29292
; Sequence 29292, Application US/10029386
; Publication No. US20030194704A1

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29292
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.79
; OTHER INFORMATION: SWISSPROT HIT: P27487, EVALUATE 4.00e-17
US-10-029-386-29292

Query Match 100.0%; Score 65; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YALWWSPNGK 10
|||||

Db 11 YALWWSPNGK 20

RESULT 4

US-09-265-606-2

; Sequence 2, Application US/09265606
; Patent No. US20020034789A1

; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789A1man D.

REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5330.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-265-606-2

Query Match 100.0%; Score 65; DB 3; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
| | | | | | | | | |
Db 210 YALWSPNGK 219

RESULT 5

US-10-177-293-136
Sequence 136, Application US/10177293
Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Baat Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.

TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-038

CURRENT APPLICATION NUMBER: US/10/177,293

CURRENT FILING DATE: 2002-06-21

PRIOR APPLICATION NUMBER: US 60/299,887

PRIOR FILING DATE: 2001-06-21

PRIOR APPLICATION NUMBER: US 60/301,572

PRIOR FILING DATE: 2001-06-27

PRIOR APPLICATION NUMBER: US 60/306,501

PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: US 60/325,002

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/362,585

PRIOR FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/xxx,xxx

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 506

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 136

LENGTH: 760

TYPE: PRT

ORGANISM: Homo sapiens

US-10-177-293-136

Query Match 100.0%; Score 65; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
| | | | | | | | | |
Db 210 YALWSPNGK 219

RESULT 6

US-10-301-822-55
Sequence 55, Application US/10301822
Publication No. US20030148410A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Berger, Allison
APPLICANT: Guillemette, Tracy L.
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: BURGART, Lawrence J.

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF COLON CANCER

FILE REFERENCE: MPM01-029P2RNM

CURRENT APPLICATION NUMBER: US/10/301,822

CURRENT FILING DATE: 2002-11-21

PRIOR APPLICATION NUMBER: US 60/339,971

PRIOR FILING DATE: 2001-12-10

PRIOR APPLICATION NUMBER: US 60/361,978

PRIOR FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/381,988

PRIOR FILING DATE: 2002-05-20

NUMBER OF SEQ ID NOS: 228

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 55

LENGTH: 760

TYPE: PRT

ORGANISM: Homo Sapiens

US-10-301-822-55

Query Match 100.0%; Score 65; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YALWSPNGK 10
| | | | | | | | | |
Db 210 YALWSPNGK 219

RESULT 7

US-10-723-860-4171

Sequence 4171, Application US/10723860

Publication No. US20040253606A1

GENERAL INFORMATION:

APPLICANT: Aziz, Natasha

APPLICANT: Ginsburg, Wendy M.

APPLICANT: Zlotnik, Albert

TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions & Methods for Screening for Soft Tissue Sarcoma Modulators

FILE REFERENCE: 05882.0193.NPUS01

CURRENT APPLICATION NUMBER: US/10/723,860

CURRENT FILING DATE: 2003-11-26

PRIOR APPLICATION NUMBER: 60/429,739

PRIOR FILING DATE: 2002-11-26

NUMBER OF SEQ ID NOS: 8393

SOFTWARE: PatentIn version 3.2

SEQ ID NO 4171

LENGTH: 760

TYPE: PRT

ORGANISM: Homo sapiens

US-10-723-860-4171

Query Match 100.0%; Score 65; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 YALWWSNPKG 10
Db      210 YALWWSNPKG 219

RESULT 8
US-10-884-070A-13
; Sequence 13, Application US/10884070A
; Publication No. US20050170368A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Cairns, Belinda
; APPLICANT: Dowd, Patrick
; APPLICANT: Frantz, Gretchen
; APPLICANT: French, Dorothy
; APPLICANT: Gonzalez, Lino
; APPLICANT: Polakis, Paul
; APPLICANT: Smith, Victoria
; APPLICANT: Wolf, Beni B.
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF TUMOR
; FILE REFERENCE: P5036R1-US
; CURRENT APPLICATION NUMBER: US/10/884,070A
; CURRENT FILING DATE: 2004-07-02
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 13
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-884-070A-13

Query Match      100.0%; Score 65; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YALWWSNPKG 10
Db      210 YALWWSNPKG 219

RESULT 9
US-10-072-012-863
; Sequence 863, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 863
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Dipeptidyl
; OTHER INFORMATION: peptidase IV (DPP IV) N-terminal region
US-10-072-012-863

Query Match      84.6%; Score 55; DB 4; Length 504;
Best Local Similarity 88.9%; Pred. No. 8;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 YALWWSNPKG 9
Db      171 YALWWSNPKG 179

RESULT 10
US-10-476-264-127
; Sequence 127, Application US/10476264
; Publication No. US20050123910A1
; GENERAL INFORMATION:
; APPLICANT: Cookson, William Osmond Charles Michael
; APPLICANT: Moffat, Miriam Fleur
; APPLICANT: Allen, Maxine
; APPLICANT: Lench, Nick
; TITLE OF INVENTION: Enzyme and SNP marker for disease
; FILE REFERENCE: 16721-002U1
; CURRENT APPLICATION NUMBER: US/10/476,264
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/GB02/01887
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: GB0110044.5
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0110046.0
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0124594.3
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: GB0124575.2
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 421
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 127
; LENGTH: 48
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-476-264-127

Query Match      81.5%; Score 53; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPNG 9
      |||||
Db      15 ALWWSPNG 22

RESULT 11
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

Query Match      81.5%; Score 53; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPNG 9
      |||||
Db      213 ALWWSPNG 220

RESULT 12
US-09-993-959-1
; Sequence 1, Application US/09993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
```

```
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-959-1

Query Match      81.5%; Score 53; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPNG 9
      |||||
Db      213 ALWWSPNG 220

RESULT 13
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match      81.5%; Score 53; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPNG 9
      |||||
Db      213 ALWWSPNG 220

RESULT 14
US-10-165-603-7
; Sequence 7, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; APPLICANT: Stevens, Anthony C.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; FILE REFERENCE: TPECH.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 766
```

TYPE: PRT
US-10-165-603-7

Query Match 81.5%; Score 53; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPNG 9

Db 213 ALWWSPNG 220

RESULT 15

US-10-423-714-6
Sequence 6, Application US/10423714
Publication No. US20030180828A1
GENERAL INFORMATION:
APPLICANT: Vanderbilt University
APPLICANT: Brown, Nancy J.
TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
FILE REFERENCE: Atty Docket No. US20030180828A1 1242/48/2
CURRENT FILING DATE: 2003-04-25
CURRENT APPLICATION NUMBER: US/10/423,714
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6

LENGTH: 766

TYPE: PRT
US-10-423-714-6

Query Match 81.5%; Score 53; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPNG 9

Db 213 ALWWSPNG 220

RESULT 16

US-10-295-027-590
Sequence 590, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: 2002-11-13
CURRENT APPLICATION NUMBER: US/10/295,027
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 922

PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 590

LENGTH: 766

TYPE: PRT
US-10-295-027-590

Query Match 81.5%; Score 53; DB 4; Length 766;

Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPNG 9

Db 213 ALWWSPNG 220

RESULT 17

US-10-295-027-922
Sequence 922, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: 2002-11-13
CURRENT APPLICATION NUMBER: US/10/295,027
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 922

```
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-922

Query Match      81.5%; Score 53; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
Db      213 ALWWSPPNG 220

RESULT 18
US-10-794-899-41
; Sequence 41, Application US/10794899
; Publication No. US20040146516A1
; GENERAL INFORMATION:
; APPLICANT: Utah Ventures
; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
; FILE REFERENCE: 27110-715
; CURRENT APPLICATION NUMBER: US/10/794,899
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-794-899-41

Query Match      81.5%; Score 53; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
Db      213 ALWWSPPNG 220

RESULT 19
US-10-952-459-18
; Sequence 18, Application US/10952459
; Publication No. US20050074805A1
; GENERAL INFORMATION:
; APPLICANT: Kochan, Jaroma Peter
; APPLICANT: Martin, Michelle Lee
; APPLICANT: Rosinski, James Andrew
; TITLE OF INVENTION: Specific Markers for Diabetes
; FILE REFERENCE: 21270US1
; CURRENT APPLICATION NUMBER: US/10/952,459
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-952-459-18

Query Match      81.5%; Score 53; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
Db      213 ALWWSPPNG 220

RESULT 20
US-10-476-264-147
; Sequence 147, Application US/10476264
```

```
; Publication No. US20050123910A1
; GENERAL INFORMATION:
; APPLICANT: Cookson, William Osmond Charles Michael
; APPLICANT: Moffat, Miriam Fleur
; APPLICANT: Allen, Maxine
; APPLICANT: Lench, Nick
; TITLE OF INVENTION: Enzyme and SNP marker for disease
; FILE REFERENCE: 16721-002US1
; CURRENT APPLICATION NUMBER: US/10/476,264
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/GB02/01887
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: GB0110044.5
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0110046.0
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0124594.3
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: GB0124575.2
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 421
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 147
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-476-264-147

Query Match      81.5%; Score 53; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
Db      213 ALWWSPPNG 220

RESULT 21
US-10-770-712-1
; Sequence 1, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSCI2.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-712-1

Query Match      81.5%; Score 53; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
Db      213 ALWWSPPNG 220

RESULT 22
US-10-631-467-565
; Sequence 565, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive f
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1003-000
```

; CURRENT APPLICATION NUMBER: US/10/631,467
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: JP 2003-077212
 ; PRIOR FILING DATE: 2003-03-20
 ; PRIOR APPLICATION NUMBER: JP 2002-229312
 ; PRIOR FILING DATE: 2002-08-06
 ; NUMBER OF SEQ ID NOS: 2086
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 565
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-631-467-565

Query Match 81.5%; Score 53; DB 5; Length 766;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||
 Db 213 ALWWSPPNG 220

RESULT 23
 US-11-041-674-6
 ; Sequence 6, Application US/11041674
 ; Publication No. US20050181468A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vanderbilt University
 ; APPLICANT: Brown, Nancy J.
 ; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING
 ; FILE REFERENCE: Acty Docket No. 1342/48/2/2/2
 ; CURRENT APPLICATION NUMBER: US/11/041,674
 ; CURRENT FILING DATE: 2005-01-24
 ; PRIOR APPLICATION NUMBER: 60/244,524
 ; PRIOR FILING DATE: 2000-10-31
 ; PRIOR APPLICATION NUMBER: 10/002,593
 ; PRIOR FILING DATE: 2001-10-31
 ; PRIOR APPLICATION NUMBER: 10/423,714
 ; PRIOR FILING DATE: 2003-4-25
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-041-674-6

Query Match 81.5%; Score 53; DB 6; Length 766;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||
 Db 213 ALWWSPPNG 220

RESULT 24
 US-10-165-603-4
 ; Sequence 4, Application US/10165603
 ; Publication No. US20030021792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roben, Paul W.
 ; APPLICANT: Stevens, Anthony C.
 ; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
 ; FILE REFERENCE: TPECH.001A
 ; CURRENT APPLICATION NUMBER: US/10/165,603
 ; CURRENT FILING DATE: 2002-06-07
 ; PRIOR APPLICATION NUMBER: 60/297,021
 ; PRIOR FILING DATE: 2001-06-08
 ; PRIOR APPLICATION NUMBER: 60/305,117

; PRIOR FILING DATE: 2001-07-12
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 767
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-165-603-4

Query Match 81.5%; Score 53; DB 4; Length 767;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||
 Db 211 ALWWSPPNG 218

RESULT 25
 US-10-794-899-38
 ; Sequence 38, Application US/10794899
 ; Publication No. US20040146516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Utah Ventures
 ; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
 ; FILE REFERENCE: 27110-715
 ; CURRENT APPLICATION NUMBER: US/10/794,899
 ; CURRENT FILING DATE: 2004-03-05
 ; NUMBER OF SEQ ID NOS: 121
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 38
 ; LENGTH: 767
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-794-899-38

Query Match 81.5%; Score 53; DB 4; Length 767;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||
 Db 211 ALWWSPPNG 218

RESULT 26
 US-10-770-712-2
 ; Sequence 2, Application US/10770712
 ; Publication No. US20050170333A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Voildani, Aristo
 ; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
 ; FILE REFERENCE: IMSCI2.008A
 ; CURRENT APPLICATION NUMBER: US/10/770,712
 ; CURRENT FILING DATE: 2004-02-03
 ; NUMBER OF SEQ ID NOS: 133
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 767
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-770-712-2

Query Match 81.5%; Score 53; DB 5; Length 767;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSPPNG 9
 |||||
 Db 211 ALWWSPPNG 218

RESULT 27


```
US-10-156-761-13020
; Sequence 13020, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13020
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13020

Query Match 75.4%; Score 49; DB 4; Length 404;
Best Local Similarity 77.8%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
Db 349 YALWVPHG 357

RESULT 28
US-10-770-712-3
; Sequence 3, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-770-712-3

Query Match 72.3%; Score 47; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSN 8
Db 207 ALWWSN 213

RESULT 29
US-10-631-467-1390
; Sequence 1390, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive P
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31

US-10-156-761-13020
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1390
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-631-467-1390

Query Match 72.3%; Score 47; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSN 8
Db 207 ALWWSN 213

RESULT 30
US-10-369-493-7006
; Sequence 7006, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7006
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-7006

Query Match 72.3%; Score 47; DB 4; Length 829;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWSNPKG 10
Db 265 AVWWSPSGR 273

RESULT 31
US-10-723-807-2
; Sequence 2, Application US/10723807
; Publication No. US20040171104A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Kimberly Brown
; APPLICANT: Michael W. Rey
; APPLICANT: Alan Klotz
; APPLICANT: Tony Byun
; TITLE OF INVENTION: Polypeptides Having Dipeptidyl
; FILE REFERENCE: 5254.200-US
; CURRENT APPLICATION NUMBER: US/10/723,807
; CURRENT FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: US/09/079,592
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 08/857,884
; PRIOR FILING DATE: 1997-05-16
```

;; PRIOR APPLICATION NUMBER: 60/062,892
;; PRIOR FILING DATE: 1997-10-20
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 771
;; TYPE: PRT
;; ORGANISM: Aspergillus oryzae
US-10-723-807-2

Query Match 70.8%; Score 46; DB 4; Length 771;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
DB 205 YALWSPDGE 214
|||||:|

RESULT 32
US-10-369-493-19475 Application US/10369493
; Sequence 19475, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19475
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19475

Query Match 69.2%; Score 45; DB 4; Length 661;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WSPNGK 10
DB 171 WSPDGG 177
|||||:

RESULT 33
US-10-369-493-2301
; Sequence 2301, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2301
; LENGTH: 758
; TYPE: PRT

;; ORGANISM: Schizosaccharomyces pombe
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(758)
;; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-2301

Query Match 69.2%; Score 45; DB 4; Length 758;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWSPNG 9
DB 178 AVWSPDG 185
|:|||||:

RESULT 34
US-10-336-597-4
; Sequence 4, Application US/10336597
; Publication No. US20030150021A1
; GENERAL INFORMATION:
; APPLICANT: KREBBERS, ENNO
; BROGLIE, KAREN E.
; PEARLSTEIN, RICHARD W.
; TITLE OF INVENTION: PLANT 4-(-GLUCANOTRANSFERASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DUPONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (VER. 7.0)
; CURRENT APPLICATION NUMBER: US/10/336,597
; FILING DATE: 03-Jan-2003
; CLASSIFICATION: <Unknown>
; APPLICATION NUMBER: US/09/403,332
; FILING DATE: 19-Oct-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: MAJARIAN, WILLIAM R.
; REGISTRATION NUMBER: 41,173
; REFERENCE/DOCKET NUMBER: BB-1101-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4926
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-336-597-4

Query Match 67.7%; Score 44; DB 4; Length 323;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
DB 90 FRLWIPKKG 99
|:|||||:

RESULT 35
US-10-156-761-9740

```
; Sequence 9740, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9740
; LENGTH: 1256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9740

Query Match      67.7%; Score 44; DB 4; Length 1256;
Best Local Similarity 85.7%; Pred. No. 6.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 LWMSPNG 9
DB      1002 LWMSPSG 1008

RESULT 36
US-10-425-115-305249
; Sequence 305249, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 305249
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_41458C.1.pap
US-10-425-115-305249

Query Match      66.9%; Score 43.5; DB 4; Length 87;
Best Local Similarity 72.7%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY      1 YALMW-SPNGK 10
DB      36 YQLWMPSPRGK 46

RESULT 37
US-10-450-763-40309
; Sequence 40309, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40309
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-40309

Query Match      66.2%; Score 43; DB 5; Length 112;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 WWSPPNGK 10
DB      103 WWPENGR 109

RESULT 38
US-10-432-198-4
; Sequence 4, Application US/10432198
; Publication No. US20040096899A1
; GENERAL INFORMATION:
; APPLICANT: AOKI, Takanori
; APPLICANT: YONEZAWA, Kayoko
; APPLICANT: FUJIMOTO, Noboru
; APPLICANT: OGAWA, Miwa
; APPLICANT: IWATA, Kazushi
; TITLE OF INVENTION: Immunoassay for Membrane-type Matrix Metalloproteinases
; FILE REFERENCE: 2003-0679A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/432,198
; CURRENT FILING DATE: 2003-10-27
; PRIOR APPLICATION NUMBER: JP 2000-352491
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-432-198-4

Query Match      66.2%; Score 43; DB 4; Length 189;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWWSPPNGK 10
DB      100 ALFWMPNGK 108

RESULT 39
US-10-432-198-5
; Sequence 5, Application US/10432198
; Publication No. US20040096899A1
; GENERAL INFORMATION:
; APPLICANT: AOKI, Takanori
; APPLICANT: YONEZAWA, Kayoko
; APPLICANT: FUJIMOTO, Noboru
; APPLICANT: OGAWA, Miwa
; APPLICANT: IWATA, Kazushi
; TITLE OF INVENTION: Immunoassay for Membrane-type Matrix Metalloproteinases
; FILE REFERENCE: 2003-0679A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/432,198
; CURRENT FILING DATE: 2003-10-27
```

; PRIOR APPLICATION NUMBER: JP 2000-352491
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-432-198-5

Query Match 66.2%; Score 43; DB 4; Length 189;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
||:|||||
Db 100 ALFWMPNGK 108

RESULT 40
US-10-432-198-6
; Sequence 6, Application US/10432198
; Publication No. US20040096899A1
; GENERAL INFORMATION:
; APPLICANT: AOKI, Takanori
; APPLICANT: YONEZAWA, Kayoko
; APPLICANT: FUJIMOTO, Noboru
; APPLICANT: OGAWA, Miwa
; APPLICANT: IWATA, Kazushi
; TITLE OF INVENTION: Immunoassay for Membrane-type Matrix Metalloproteinases
; FILE REFERENCE: 2003-0679A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/432,198
; CURRENT FILING DATE: 2003-10-27
; PRIOR APPLICATION NUMBER: JP 2000-352491
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-432-198-6

Query Match 66.2%; Score 43; DB 4; Length 189;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
||:|||||
Db 100 ALFWMPNGK 108

RESULT 41
US-10-432-198-7
; Sequence 7, Application US/10432198
; Publication No. US20040096899A1
; GENERAL INFORMATION:
; APPLICANT: AOKI, Takanori
; APPLICANT: YONEZAWA, Kayoko
; APPLICANT: FUJIMOTO, Noboru
; APPLICANT: OGAWA, Miwa
; APPLICANT: IWATA, Kazushi
; TITLE OF INVENTION: Immunoassay for Membrane-type Matrix Metalloproteinases
; FILE REFERENCE: 2003-0679A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/432,198
; CURRENT FILING DATE: 2003-10-27
; PRIOR APPLICATION NUMBER: JP 2000-352491
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 189
; TYPE: PRT

; ORGANISM: Oryctolagus cuniculus
US-10-432-198-7

Query Match 66.2%; Score 43; DB 4; Length 189;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
||:|||||
Db 100 ALFWMPNGK 108

RESULT 42
US-10-276-272A-7
; Sequence 7, Application US/10276272A
; Publication No. US20040120954A1
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: OBATA, Ken-ichi
; TITLE OF INVENTION: REGULATION OF MT1-MMP ACTIVITY
; FILE REFERENCE: 2002-1357A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/276,272A
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: PCT/JP01/04166
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: JP2000/147739
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 20
; SEQ ID NO 7
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MT1PEX
US-10-276-272A-7

Query Match 66.2%; Score 43; DB 4; Length 218;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPPNGK 10
||:|||||
Db 101 ALFWMPNGK 109

RESULT 43
US-10-424-599-225890
; Sequence 225890, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 225890
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46009C.1.pap
US-10-424-599-225890

Query Match 66.2%; Score 43; DB 4; Length 280;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 YALWWSPPNGK 10

```
Db      238 YAVSWSPDGK 247  
||: |||:||  
  
RESULT 44  
US-10-276-272A-10  
; Sequence 10, Application US/10276272A  
; Publication No. US20040120954A1  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: OBATA, Ken-ichi  
; TITLE OF INVENTION: REGULATION OF MT1-MMP ACTIVITY  
; FILE REFERENCE: 2002-1357A/WMC/01332  
; CURRENT FILING DATE: 2003-12-02  
; PRIOR FILING DATE: 2003-12-02  
; PRIOR APPLICATION NUMBER: PCT/JPO1/04166  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: JP2000/147739  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 20  
; SEQ ID NO 10  
; LENGTH: 381  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MT1PEX  
US-10-276-272A-10  
  
Query Match          66.2%; Score 43; DB 4; Length 381;  
Best Local Similarity 77.8%; Pred.No. 3.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0;  
  
Qy      2 ALMWSPNGK 10  
||: |||:||  
Db      217 ALFWMPNGK 225  
  
RESULT 45  
US-10-276-272A-4  
; Sequence 4, Application US/10276272A  
; Publication No. US20040120954A1  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: OBATA, Ken-ichi  
; TITLE OF INVENTION: REGULATION OF MT1-MMP ACTIVITY  
; FILE REFERENCE: 2002-1357A/WMC/01332  
; CURRENT APPLICATION NUMBER: US/10/276, 272A  
; CURRENT FILING DATE: 2003-12-02  
; PRIOR FILING DATE: 2003-12-02  
; PRIOR APPLICATION NUMBER: PCT/JPO1/04166  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: JP2000/147739  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 20  
; SEQ ID NO 4  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: MT1PEX-F  
US-10-276-272A-4  
  
Query Match          66.2%; Score 43; DB 4; Length 389;  
Best Local Similarity 77.8%; Pred.No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy      2 ALMWSPNGK 10  
||: |||:||  
Db      225 ALFWMPNGK 233  
  
RESULT 46  
US-10-276-272A-6  
; Sequence 6, Application US/10276272A
```

Mon Mar 13 10:54:36 2006

```
Db      307 ALFWMPNGK 315
||:| ||||
RESULT 48
US-09-916-658-4
; Sequence 4, Application US/09916658
; Patent No. US2002025510A1
; GENERAL INFORMATION:
; APPLICANT: Strongin, Alex Y.
; APPLICANT: Deryugina, Elena I.
; TITLE OF INVENTION: Screening Methods Based On
; TITLE OF INVENTION: Superactivated Alpha v Beta 3 Integrin
; FILE REFERENCE: P-LJ 4811
; CURRENT APPLICATION NUMBER: US/09/916,658
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: US 60/220,706
; PRIOR FILING DATE: 2000-07-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-916-658-4

Query Match      66.2%; Score 43; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWSPNGK 10
||:| ||||
Db      418 ALFWMPNGK 426

RESULT 49
US-09-391-104-28
; Sequence 28, Application US/093911104
; Publication No. US20020031817A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Falduto, Michael T.
; APPLICANT: Magnuson, Scott R.
; APPLICANT: Morgan, Douglas W.
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEINASE GENE,
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS
; FILE REFERENCE: 6073.US.P1
; CURRENT APPLICATION NUMBER: US/09/391,104
; CURRENT FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/814,394
; PRIOR FILING DATE: 1997-03-11
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-391-104-28

Query Match      66.2%; Score 43; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWSPNGK 10
||:| ||||
Db      418 ALFWMPNGK 426

RESULT 50
US-09-801-196-27
; Sequence 27, Application US/09801196
; Patent No. US20020037827A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Kai
; APPLICANT: Smith, Ryan
; APPLICANT: Fajardo, Mark
; APPLICANT: Moss, Patrick
; TITLE OF INVENTION: A NOVEL MATRIX METALLOPROTEINASE (MMP-25)
; TITLE OF INVENTION: EXPRESSED IN SKIN CELLS
; FILE REFERENCE: 240083.509
; CURRENT APPLICATION NUMBER: US/09/801,196
; CURRENT FILING DATE: 2001-03-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-801-196-27

Query Match      66.2%; Score 43; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWSPNGK 10
||:| ||||
Db      418 ALFWMPNGK 426

RESULT 51
US-09-919-497-84
; Sequence 84, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-84

Query Match      66.2%; Score 43; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWSPNGK 10
||:| ||||
Db      418 ALFWMPNGK 426

RESULT 52
US-09-916-849A-2
; Sequence 2, Application US/09916849A
; Publication No. US20030086934A1
; GENERAL INFORMATION:
; APPLICANT: Bostein, et al.
; TITLE OF INVENTION: Basal Markers in Breast Cancer and Related Reagents
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 2002850-0024
; CURRENT APPLICATION NUMBER: US/09/916,849A
; CURRENT FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-916-849A-2
```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Matrix
; OTHER INFORMATION: Metalloproteinase
US-09-916-849A-2

Query Match 66.2%; Score 43; DB 3; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
DB 418 ALFWMPNGK 426

RESULT 53

US-10-133-797-6
; Sequence 6, Application US/10133797
; Publication No. US20030109021A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Shujian
; APPLICANT: Chen, Jian
; APPLICANT: Feder, John
; APPLICANT: Lee, Liana
; APPLICANT: Krystek, Stanley
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEINASE HIGHLY
; FILE REFERENCE: D0141NP
; CURRENT APPLICATION NUMBER: US/10/133,797
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/286,764
; PRIOR FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-133-797-6

Query Match 66.2%; Score 43; DB 4; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
DB 418 ALFWMPNGK 426

RESULT 54

US-10-131-985-43
; Sequence 43, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-131-985-43

Query Match 66.2%; Score 43; DB 4; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
DB 418 ALFWMPNGK 426

RESULT 55

US-10-411-010-26
; Sequence 26, Application US/10411010
; Publication No. US20030228319A1
; GENERAL INFORMATION:
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: PAUL POLAKIS
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P502381-US
; CURRENT APPLICATION NUMBER: US/10/411,010
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/378,885
; PRIOR FILING DATE: 2002-05-05
; PRIOR APPLICATION NUMBER: US 60/373,160
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 26
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-411-010-26

Query Match 66.2%; Score 43; DB 4; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWWSPNGK 10
||:|||||
DB 418 ALFWMPNGK 426

RESULT 56

US-10-447-315-9
; Sequence 9, Application US/10447315
; Publication No. US20040071687A1
; GENERAL INFORMATION:
; APPLICANT: Rafii, Shahin
; APPLICANT: Heissig, Beate
; APPLICANT: Hattori, Koichi
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Adult Stem Cell Recruitment
; FILE REFERENCE: 1676.006US1
; CURRENT APPLICATION NUMBER: US/10/447,315
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: US 60/383,658
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-447-315-9

Query Match 66.2%; Score 43; DB 4; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 57
US-10-276-272A-19
; Sequence 19, Application US/10276272A
; Publication No. US20040120954A1
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: OBATA, Ken-ichi
; TITLE OF INVENTION: REGULATION OF MT1-MMP ACTIVITY
; FILE REFERENCE: 2002-1357A/WMC/01332
; CURRENT APPLICATION NUMBER: US/10/276,272A
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: PCT/JP01/04166
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: JP2000/147739
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 20
; SEQ ID NO 19
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MT1-MMP
US-10-276-272A-19

Query Match 66.2%; Score 43; DB 4; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 58
US-10-901-417-43
; Sequence 43, Application US/10901417
; Publication No. US20050026836A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/901,417
; CURRENT FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US/10/131,985
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-901-417-43

Query Match 66.2%; Score 43; DB 5; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 59
US-10-953-264-26
; Sequence 26, Application US/10953264
; Publication No. US20050042216A1
; GENERAL INFORMATION:
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: PAUL POLAKIS
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5023R1-US
; CURRENT APPLICATION NUMBER: US/10/953,264
; CURRENT FILING DATE: 2004-09-29
; PRIOR APPLICATION NUMBER: US/10/411,010
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/378,885
; PRIOR FILING DATE: 2002-05-05
; PRIOR APPLICATION NUMBER: US 60/373,160
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 26
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-953-264-26

Query Match 66.2%; Score 43; DB 5; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSNGK 10
||:|||||
Db 418 ALFWMPNGK 426

RESULT 60
US-10-773-446-100
; Sequence 100, Application US/10773446
; Publication No. US20050176662A1
; GENERAL INFORMATION:
; APPLICANT: INANA, GEORGE
; APPLICANT: McLAREN, MARGARET
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING AND TREATING RETINAL
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: 39532-192229
; CURRENT APPLICATION NUMBER: US/10/773,446
; CURRENT FILING DATE: 2004-02-09
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 100
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-773-446-100

Query Match 66.2%; Score 43; DB 5; Length 582;
Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSNGK 10
||:|||||
Db 418 ALFWMPNGK 426

Search completed: March 11, 2006, 12:32:25
Job time : 84.4211 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:27:37 ; Search time 9.21053 Seconds
(without alignments)
30.221 Million cell updates/sec

Title: US-10-774-242A-4

Perfect score: 65

Sequence: 1 VALMWSPNGK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications AA New:**

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep:**
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep:**
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep:**
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep:**
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep:**
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep:**
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep:**
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	760	7	US-11-186-284-55
2	53	81.5	782	7	US-11-116-939-13
3	53	81.5	766	6	US-10-522-789-2
4	53	81.5	766	6	US-10-501-035-234
5	43	66.2	582	7	US-11-090-439-58
6	43	66.2	582	7	US-11-169-041-130
7	43	66.2	582	7	US-11-200-822-2
8	41	63.1	984	7	US-11-024-959-477
9	40	61.5	475	7	US-11-087-099-7087
10	39	60.0	2335	6	US-10-821-234-1610
11	39	60.0	174	7	US-11-009-658-16
12	39	60.0	255	7	US-11-009-658-42
13	39	60.0	366	7	US-11-024-959-470
14	39	60.0	1234	6	US-10-995-561-870
15	39	60.0	1365	6	US-10-995-561-867
16	39	60.0	1366	6	US-10-995-561-868
17	39	60.0	1411	6	US-10-995-561-869
18	38	58.5	104	7	US-11-072-512-2562
19	38	58.5	228	7	US-11-098-686-10864
20	37	56.9	127	7	US-11-116-144-91
21	37	56.9	129	7	US-11-116-144-90
22	37	56.9	261	7	US-11-096-568A-21612
23	37	56.9	271	7	US-11-096-568A-21611
24	37	56.9	272	7	US-11-096-568A-10627
25	37	56.9	327	7	US-11-096-568A-21610

26	37	56.9	328	7	US-11-096-568A-10626	Sequence 10626, A
27	37	56.9	370	7	US-11-096-568A-10625	Sequence 10625, A
28	37	56.9	606	7	US-11-096-568A-17843	Sequence 17843, A
29	37	56.9	607	7	US-11-024-959-381	Sequence 381, Appl
30	36	55.4	9	7	US-11-107-096-51	Sequence 51, Appl
31	36	55.4	209	7	US-11-096-568A-3423	Sequence 3423, Ap
32	36	55.4	929	6	US-10-467-657-5556	Sequence 5656, Ap
33	35.5	54.6	131	7	US-11-049-536-84	Sequence 84, Appl
34	35.5	54.6	318	7	US-11-112-882-32	Sequence 32, Appl
35	35	53.8	119	6	US-10-925-366A-352	Sequence 352, App
36	35	53.8	129	6	US-10-993-543-180	Sequence 180, App
37	35	53.8	282	7	US-11-098-686-10681	Sequence 10681, A
38	35	53.8	303	6	US-10-063-703-166	Sequence 166, App
39	35	53.8	303	7	US-11-102-240-186	Sequence 162, Appl
40	35	53.8	311	7	US-11-241-056-12	Sequence 12, Appl
41	35	53.8	441	6	US-11-098-686-10683	Sequence 10683, A
42	35	53.8	445	6	US-10-517-939-368	Sequence 368, App
43	35	53.8	463	7	US-11-024-959-507	Sequence 507, App
44	35	53.8	479	6	US-10-517-939-250	Sequence 250, App
45	35	53.8	484	7	US-11-072-512-3591	Sequence 3591, Ap
46	35	53.8	508	7	US-11-087-099-9085	Sequence 9085, Ap
47	35	53.8	547	7	US-11-087-099-10512	Sequence 10512, A
48	35	53.8	607	7	US-11-080-991-88	Sequence 88, Appl
49	35	53.8	608	7	US-11-024-959-481	Sequence 481, App
50	35	53.8	791	7	US-11-072-512-3453	Sequence 3453, Ap
51	34	52.3	36	6	US-10-613-744-41	Sequence 41, Appl
52	34	52.3	40	6	US-10-613-744-23	Sequence 23, Appl
53	34	52.3	58	6	US-10-613-744-3	Sequence 3, Appl
54	34	52.3	58	6	US-10-613-744-25	Sequence 25, Appl
55	34	52.3	91	7	US-11-096-568A-27334	Sequence 27334, A
56	34	52.3	134	7	US-11-087-099-4615	Sequence 4615, Ap
57	34	52.3	156	7	US-11-096-568A-15758	Sequence 15758, A
58	34	52.3	160	6	US-10-613-744-1	Sequence 1, Appl
59	34	52.3	160	6	US-10-613-744-16	Sequence 16, Appl
60	34	52.3	181	7	US-11-072-512-2677	Sequence 2677, Ap
61	34	52.3	225	7	US-11-087-099-5166	Sequence 5166, Ap
62	34	52.3	320	7	US-11-087-099-747	Sequence 747, App
63	34	52.3	322	7	US-11-096-568A-13003	Sequence 13003, A
64	34	52.3	342	7	US-11-096-568A-7217	Sequence 7217, Ap
65	34	52.3	343	7	US-11-096-568A-7216	Sequence 7216, Ap
66	34	52.3	352	7	US-11-108-528-20	Sequence 20, Appl
67	34	52.3	352	7	US-11-108-528-22	Sequence 22, Appl
68	34	52.3	355	7	US-11-108-528-16	Sequence 16, Appl
69	34	52.3	355	7	US-11-108-528-18	Sequence 18, Appl
70	34	52.3	387	7	US-11-087-099-12140	Sequence 12140, A
71	34	52.3	430	7	US-11-096-568A-2421	Sequence 2421, Ap
72	34	52.3	439	6	US-10-453-372-270	Sequence 270, App
73	34	52.3	445	7	US-11-072-512-2209	Sequence 2209, Ap
74	34	52.3	452	7	US-11-087-099-11368	Sequence 11368, A
75	34	52.3	466	7	US-11-096-568A-2420	Sequence 2420, Ap
76	34	52.3	475	7	US-11-024-959-354	Sequence 354, App
77	34	52.3	513	7	US-11-087-099-1119	Sequence 1119, Ap
78	34	52.3	529	7	US-11-096-568A-2419	Sequence 2419, Ap
79	34	52.3	529	7	US-11-096-568A-2422	Sequence 2422, Ap
80	34	52.3	577	6	US-10-501-035-240	Sequence 240, App
81	34	52.3	629	6	US-10-453-372-268	Sequence 268, App
82	34	52.3	742	6	US-10-453-372-264	Sequence 264, App
83	34	52.3	1330	6	US-10-453-372-260	Sequence 260, App
84	34	52.3	1542	6	US-10-453-372-258	Sequence 258, App
85	34	52.3	1542	6	US-10-453-372-266	Sequence 266, App
86	34	52.3	1542	6	US-10-453-372-280	Sequence 280, App
87	34	52.3	1542	6	US-10-453-372-280	Sequence 280, Appl
88	33	50.8	112	7	US-11-031-206-88	Sequence 88, Appl
89	33	50.8	128	7	US-11-116-144-148	Sequence 148, App
90	33	50.8	134	7	US-11-096-568A-4045	Sequence 4045, Ap
91	33	50.8	160	7	US-11-096-568A-17266	Sequence 17266, A
92	33	50.8	176	7	US-11-096-568A-26064	Sequence 26064, A
93	33	50.8	185	7	US-11-096-568A-26064	Sequence 90, Appl
94	33	50.8	190	7	US-11-031-206-90	Sequence 26063, A
95	33	50.8	206	7	US-11-096-568A-26063	Sequence 5057, Ap
96	33	50.8	246	7	US-11-096-568A-5057	Sequence 16067, A
97	33	50.8	258	7	US-11-096-568A-16067	
98	33	50.8				

99	33	50.8	258	7	US-11-096-568A-19499	Sequence 19499, A	172	32.5	50.0	594	7	US-11-072-512-3281	Sequence 3281, Ap
100	33	50.8	264	7	US-11-096-568A-16066	Sequence 16066, A	173	32	49.2	18	6	US-10-895-064-39	Sequence 39, Appl
101	33	50.8	264	7	US-11-096-568A-19498	Sequence 19498, A	174	32	49.2	18	7	US-11-129-741-39	Sequence 39, Appl
102	33	50.8	269	6	US-10-972-587-16	Sequence 16, Appl	175	32	49.2	18	7	US-11-129-741-2975	Sequence 2975, Ap
103	33	50.8	274	7	US-11-096-568A-23575	Sequence 23575, A	176	32	49.2	121	7	US-11-098-686-10845	Sequence 10845, A
104	33	50.8	277	7	US-11-096-568A-5056	Sequence 5056, Ap	177	32	49.2	124	7	US-11-096-568A-13115	Sequence 13115, A
105	33	50.8	281	7	US-11-096-568A-5055	Sequence 5055, Ap	178	32	49.2	129	7	US-11-116-144-114	Sequence 114, App
106	33	50.8	283	7	US-11-096-568A-11239	Sequence 11239, A	179	32	49.2	130	7	US-11-096-568A-20611	Sequence 20611, A
107	33	50.8	292	7	US-11-096-568A-16065	Sequence 16065, A	180	32	49.2	140	6	US-10-980-388-93	Sequence 93, Appl
108	33	50.8	297	7	US-11-096-568A-19497	Sequence 19497, A	181	32	49.2	150	7	US-11-096-568A-4971	Sequence 4971, Ap
109	33	50.8	315	6	US-10-467-657-7848	Sequence 7848, Ap	182	32	49.2	159	7	US-11-174-150-31	Sequence 31, Appl
110	33	50.8	324	7	US-11-087-099-3409	Sequence 3409, Ap	183	32	49.2	166	7	US-11-087-099-3091	Sequence 3091, Ap
111	33	50.8	331	7	US-11-096-568A-22525	Sequence 22525, A	184	32	49.2	210	7	US-11-096-568A-4970	Sequence 4970, Ap
112	33	50.8	336	7	US-11-096-568A-22524	Sequence 22524, A	185	32	49.2	212	7	US-11-096-568A-4969	Sequence 4969, Ap
113	33	50.8	350	7	US-11-096-568A-22523	Sequence 22523, A	186	32	49.2	223	7	US-11-096-568A-25983	Sequence 25983, A
114	33	50.8	353	7	US-11-206-587-11	Sequence 11, Appl	187	32	49.2	241	6	US-11-170-653-35	Sequence 35, Appl
115	33	50.8	360	7	US-11-096-568A-11238	Sequence 11238, A	188	32	49.2	243	6	US-10-467-657-5538	Sequence 5538, Ap
116	33	50.8	376	7	US-11-087-099-10463	Sequence 10463, A	189	32	49.2	286	6	US-10-858-730-199	Sequence 199, App
117	33	50.8	400	7	US-11-096-568A-7537	Sequence 7537, Ap	190	32	49.2	288	7	US-11-087-099-3871	Sequence 3871, Ap
118	33	50.8	405	6	US-10-821-234-1357	Sequence 1357, Ap	191	32	49.2	313	6	US-10-467-657-3982	Sequence 3982, Ap
119	33	50.8	406	6	US-10-330-773-602	Sequence 602, App	192	32	49.2	313	6	US-10-467-657-8140	Sequence 8140, Ap
120	33	50.8	432	7	US-11-087-099-4508	Sequence 4508, Ap	193	32	49.2	313	7	US-11-052-554A-247	Sequence 247, App
121	33	50.8	432	7	US-11-096-568A-7536	Sequence 7536, Ap	194	32	49.2	320	6	US-10-467-657-424	Sequence 424, App
122	33	50.8	434	7	US-11-096-568A-11237	Sequence 11237, A	195	32	49.2	320	7	US-11-052-554A-239	Sequence 239, App
123	33	50.8	439	7	US-11-087-099-2654	Sequence 2654, Ap	196	32	49.2	332	7	US-11-096-568A-23521	Sequence 23521, A
124	33	50.8	442	7	US-11-087-099-3716	Sequence 3716, Ap	197	32	49.2	356	7	US-11-074-176-200	Sequence 200, App
125	33	50.8	446	7	US-11-087-099-3804	Sequence 3804, Ap	198	32	49.2	357	7	US-11-096-568A-3231	Sequence 3231, Ap
126	33	50.8	447	7	US-11-087-099-8120	Sequence 8120, Ap	199	32	49.2	387	7	US-11-096-568A-21523	Sequence 21523, A
127	33	50.8	450	7	US-11-087-099-4313	Sequence 4313, Ap	200	32	49.2	387	7	US-11-096-568A-22707	Sequence 22707, A
128	33	50.8	450	7	US-11-087-099-6129	Sequence 6129, Ap	201	32	49.2	391	7	US-11-096-568A-3230	Sequence 3230, Ap
129	33	50.8	452	6	US-10-878-556A-127	Sequence 127, App	202	32	49.2	391	7	US-11-096-568A-3232	Sequence 3232, Ap
130	33	50.8	455	7	US-11-098-686-10530	Sequence 10530, A	203	32	49.2	394	7	US-11-096-568A-23520	Sequence 23520, A
131	33	50.8	458	6	US-10-932-182A-3	Sequence 3, Appl	204	32	49.2	398	6	US-10-821-234-1583	Sequence 1583, Ap
132	33	50.8	458	6	US-10-932-182A-3	Sequence 3, Appl	205	32	49.2	400	7	US-11-087-099-419	Sequence 419, App
133	33	50.8	469	7	US-11-087-099-2731	Sequence 2731, Ap	206	32	49.2	414	7	US-11-087-099-9722	Sequence 9722, Ap
134	33	50.8	474	7	US-11-024-953-383	Sequence 383, App	207	32	49.2	435	7	US-11-087-099-8036	Sequence 8036, Ap
135	33	50.8	517	7	US-11-087-099-5778	Sequence 5778, Ap	208	32	49.2	442	7	US-11-096-568A-23519	Sequence 23519, A
136	33	50.8	519	7	US-11-087-099-4494	Sequence 4494, Ap	209	32	49.2	462	7	US-11-096-568A-3229	Sequence 3229, Ap
137	33	50.8	520	7	US-11-096-568A-29664	Sequence 29664, A	210	32	49.2	478	6	US-10-821-234-915	Sequence 915, App
138	33	50.8	540	6	US-10-858-730-293	Sequence 293, App	211	32	49.2	480	7	US-11-098-686-10650	Sequence 10650, A
139	33	50.8	540	7	US-11-235-254-2	Sequence 2, Appl	212	32	49.2	486	7	US-11-096-568A-20847	Sequence 20847, A
140	33	50.8	575	7	US-11-072-512-2221	Sequence 2221, Ap	213	32	49.2	496	7	US-11-096-568A-9251	Sequence 9251, Ap
141	33	50.8	693	6	US-10-131-826A-406	Sequence 406, App	214	32	49.2	509	7	US-11-096-568A-20846	Sequence 20846, A
142	33	50.8	693	6	US-10-973-1158-406	Sequence 406, App	215	32	49.2	518	7	US-11-096-568A-24612	Sequence 24612, A
143	33	50.8	778	6	US-10-467-657-1886	Sequence 1886, Ap	216	32	49.2	520	7	US-11-087-099-10337	Sequence 10337, A
144	33	50.8	900	7	US-11-144-987-4	Sequence 4, Appl	217	32	49.2	526	7	US-11-087-099-9483	Sequence 9483, Ap
145	33	50.8	900	7	US-11-144-987-10	Sequence 10, Appl	218	32	49.2	527	7	US-11-096-568A-13109	Sequence 13109, A
146	33	50.8	900	7	US-11-205-935-4	Sequence 4, Appl	219	32	49.2	530	7	US-11-096-568A-24611	Sequence 24611, A
147	33	50.8	900	7	US-11-205-935-10	Sequence 10, Appl	220	32	49.2	532	7	US-11-096-568A-9250	Sequence 9250, Ap
148	33	50.8	902	7	US-11-144-987-6	Sequence 6, Appl	221	32	49.2	572	7	US-11-087-099-4593	Sequence 4593, Ap
149	33	50.8	902	7	US-11-144-987-8	Sequence 8, Appl	222	32	49.2	572	7	US-11-087-099-7113	Sequence 7113, Ap
150	33	50.8	902	7	US-11-144-987-12	Sequence 12, Appl	223	32	49.2	576	7	US-11-096-568A-20845	Sequence 20845, A
151	33	50.8	902	7	US-11-144-987-14	Sequence 14, Appl	224	32	49.2	586	7	US-11-096-568A-24610	Sequence 24610, A
152	33	50.8	902	7	US-11-205-935-6	Sequence 6, Appl	225	32	49.2	590	7	US-11-098-686-10762	Sequence 10762, A
153	33	50.8	902	7	US-11-205-935-8	Sequence 8, Appl	226	32	49.2	608	7	US-11-096-568A-9249	Sequence 9249, Ap
154	33	50.8	902	7	US-11-205-935-12	Sequence 12, Appl	227	32	49.2	626	7	US-11-096-568A-9252	Sequence 9252, Ap
155	33	50.8	902	7	US-11-205-935-14	Sequence 14, Appl	228	32	49.2	622	6	US-10-467-657-1346	Sequence 1346, Ap
156	33	50.8	910	7	US-11-144-987-2	Sequence 2, Appl	229	32	49.2	626	7	US-11-010-748A-1	Sequence 1, Appl
157	33	50.8	910	7	US-11-205-935-2	Sequence 2, Appl	230	32	49.2	642	6	US-10-467-657-5900	Sequence 5900, Ap
158	33	50.8	915	7	US-11-144-987-16	Sequence 16, Appl	231	32	49.2	642	7	US-11-098-686-10457	Sequence 10457, A
159	33	50.8	915	7	US-11-144-987-22	Sequence 22, Appl	232	32	49.2	649	7	US-11-232-406A-12	Sequence 12, Appl
160	33	50.8	915	7	US-11-205-935-16	Sequence 16, Appl	233	32	49.2	751	7	US-11-012-762-26	Sequence 26, Appl
161	33	50.8	915	7	US-11-205-935-22	Sequence 22, Appl	234	32	49.2	814	6	US-10-878-556A-161	Sequence 161, App
162	33	50.8	917	7	US-11-144-987-18	Sequence 18, Appl	235	32	49.2	814	6	US-10-538-471-4	Sequence 4, Appl
163	33	50.8	917	7	US-11-144-987-20	Sequence 20, Appl	236	32	49.2	1073	6	US-11-098-686-10413	Sequence 2, Appl
164	33	50.8	917	7	US-11-144-987-24	Sequence 24, Appl	237	32	49.2	1625	6	US-10-757-832-2	Sequence 276, App
165	33	50.8	917	7	US-11-144-987-26	Sequence 26, Appl	238	32	49.2	2036	7	US-11-124-368A-276	Sequence 280, App
166	33	50.8	917	7	US-11-205-935-18	Sequence 18, Appl	239	32	49.2	2036	7	US-11-124-368A-280	Sequence 281, App
167	33	50.8	917	7	US-11-205-935-20	Sequence 20, Appl	240	32	49.2	2044	7	US-11-124-368A-281	Sequence 278, App
168	33	50.8	917	7	US-11-205-935-24	Sequence 24, Appl	241	32	49.2	2144	7	US-11-124-368A-278	Sequence 277, App
169	33	50.8	1006	6	US-11-205-935-26	Sequence 26, Appl	242	32	49.2	31.5	7	US-11-124-368A-277	Sequence 277, App
170	33	50.8	1562	7	US-10-467-657-8400	Sequence 8400, Ap	243	31.5	48.5	119	7	US-11-049-536-570	Sequence 570, App
171	33	50.8	1562	7	US-11-052-554A-211	Sequence 211, App	244	31.5	48.5	351	7	US-11-024-959-344	Sequence 344, App

245	31.5	48.5	614	7	US-11-087-099-9601	Sequence 9681, Ap	318	31	47.7	322	7	US-11-096-568A-21846	Sequence 21846, A
246	31.5	48.5	640	7	US-11-096-568A-7937	Sequence 7937, Ap	319	31	47.7	323	7	US-11-096-568A-16928	Sequence 16928, A
247	31.5	48.5	645	7	US-11-096-568A-7938	Sequence 7938, Ap	320	31	47.7	325	7	US-11-096-568A-33428	Sequence 33428, A
248	31.5	48.5	672	7	US-11-096-568A-7936	Sequence 7936, Ap	321	31	47.7	337	6	US-10-520-820-6	Sequence 6, Appl
249	31.5	48.5	686	7	US-11-096-568A-30395	Sequence 30395, A	322	31	47.7	347	7	US-11-096-568A-2546	Sequence 2546, Ap
250	31.5	48.5	754	7	US-11-096-568A-30394	Sequence 30394, A	323	31	47.7	351	7	US-11-087-099-11823	Sequence 11823, A
251	31	47.7	26	6	US-10-939-890-295	Sequence 295, App	324	31	47.7	355	7	US-11-096-568A-33656	Sequence 33656, A
252	31	47.7	100	7	US-11-096-568A-24975	Sequence 24975, A	325	31	47.7	361	7	US-11-096-568A-33655	Sequence 33655, A
253	31	47.7	107	7	US-11-096-568A-1366	Sequence 1366, Ap	326	31	47.7	367	6	US-10-528-031-3	Sequence 3, Appl
254	31	47.7	118	7	US-11-049-536-350	Sequence 350, App	327	31	47.7	367	7	US-11-096-568A-19919	Sequence 19919, A
255	31	47.7	120	7	US-11-049-536-482	Sequence 482, App	328	31	47.7	368	7	US-11-096-568A-23501	Sequence 23501, A
256	31	47.7	126	7	US-11-064-174-146	Sequence 146, App	329	31	47.7	369	7	US-11-096-568A-33654	Sequence 33654, A
257	31	47.7	129	7	US-11-116-144-93	Sequence 93, Appl	330	31	47.7	374	7	US-11-096-568A-33427	Sequence 33427, A
258	31	47.7	129	7	US-11-116-144-98	Sequence 98, Appl	331	31	47.7	383	7	US-11-096-568A-2545	Sequence 2545, Ap
259	31	47.7	129	7	US-11-116-144-106	Sequence 106, App	332	31	47.7	390	7	US-11-096-568A-16827	Sequence 16827, A
260	31	47.7	129	7	US-11-116-144-108	Sequence 108, App	333	31	47.7	399	7	US-11-024-959-489	Sequence 489, App
261	31	47.7	129	7	US-11-116-144-109	Sequence 109, App	334	31	47.7	409	7	US-11-096-568A-2544	Sequence 2544, Ap
262	31	47.7	129	7	US-11-116-144-110	Sequence 110, App	335	31	47.7	419	7	US-11-087-099-10606	Sequence 10606, A
263	31	47.7	129	7	US-11-116-144-111	Sequence 111, App	336	31	47.7	423	7	US-11-000-463-375	Sequence 375, App
264	31	47.7	129	7	US-11-116-144-112	Sequence 112, App	337	31	47.7	449	7	US-11-133-360-12	Sequence 12, Appl
265	31	47.7	129	7	US-11-116-144-113	Sequence 113, App	338	31	47.7	449	7	US-11-133-360-12	Sequence 12, Appl
266	31	47.7	129	7	US-11-116-144-115	Sequence 115, App	339	31	47.7	450	7	US-11-133-360-16	Sequence 16, Appl
267	31	47.7	130	7	US-11-116-144-142	Sequence 142, App	340	31	47.7	450	7	US-11-133-360-18	Sequence 18, Appl
268	31	47.7	130	7	US-11-116-144-143	Sequence 143, App	341	31	47.7	450	7	US-11-133-360-16	Sequence 16, Appl
269	31	47.7	131	7	US-11-096-568A-18948	Sequence 18948, A	342	31	47.7	455	7	US-11-133-360-18	Sequence 18, Appl
270	31	47.7	137	7	US-11-096-568A-1365	Sequence 1365, Ap	343	31	47.7	457	7	US-11-087-099-9561	Sequence 9561, Ap
271	31	47.7	161	7	US-11-096-568A-1364	Sequence 1364, Ap	344	31	47.7	460	7	US-11-096-568A-26145	Sequence 26145, A
272	31	47.7	169	7	US-11-096-568A-18947	Sequence 18947, A	345	31	47.7	463	7	US-11-087-099-2262	Sequence 2262, Ap
273	31	47.7	171	7	US-11-087-099-7690	Sequence 7690, Ap	346	31	47.7	463	7	US-11-087-099-3792	Sequence 3792, Ap
274	31	47.7	177	7	US-11-234-424-18	Sequence 18, Appl	347	31	47.7	464	7	US-11-024-959-488	Sequence 488, App
275	31	47.7	180	7	US-11-096-568A-11708	Sequence 11708, A	348	31	47.7	464	7	US-11-096-568A-26144	Sequence 26144, A
276	31	47.7	185	7	US-11-170-653-1	Sequence 1, Appl	349	31	47.7	469	7	US-11-133-360-14	Sequence 14, Appl
277	31	47.7	191	7	US-11-096-568A-11707	Sequence 11707, A	350	31	47.7	469	7	US-11-133-360-14	Sequence 14, Appl
278	31	47.7	193	7	US-11-096-568A-4423	Sequence 4423, Ap	351	31	47.7	472	7	US-11-096-568A-26143	Sequence 26143, A
279	31	47.7	194	7	US-11-096-568A-20191	Sequence 20191, A	352	31	47.7	524	6	US-10-878-556A-152	Sequence 152, App
280	31	47.7	206	6	US-10-467-657-524	Sequence 524, App	353	31	47.7	532	7	US-11-143-980-41	Sequence 41, Appl
281	31	47.7	213	6	US-10-517-939-188	Sequence 188, App	354	31	47.7	532	7	US-11-143-980-41	Sequence 41, Appl
282	31	47.7	213	6	US-10-517-939-224	Sequence 224, App	355	31	47.7	551	7	US-11-087-099-8478	Sequence 8478, Ap
283	31	47.7	213	6	US-10-517-939-230	Sequence 230, App	356	31	47.7	570	7	US-11-072-512-3842	Sequence 3842, Ap
284	31	47.7	213	6	US-10-517-939-302	Sequence 302, App	357	31	47.7	574	7	US-11-087-099-7415	Sequence 7415, Ap
285	31	47.7	213	7	US-11-170-653-16	Sequence 16, Appl	358	31	47.7	586	7	US-11-072-512-3411	Sequence 3411, Ap
286	31	47.7	221	7	US-11-170-653-17	Sequence 17, Appl	359	31	47.7	610	7	US-11-142-700-16	Sequence 16, Appl
287	31	47.7	221	7	US-11-170-653-20	Sequence 20, Appl	360	31	47.7	637	7	US-11-179-597-1	Sequence 1, Appl
288	31	47.7	222	6	US-10-517-939-168	Sequence 168, App	361	31	47.7	659	6	US-10-649-591-17	Sequence 17, Appl
289	31	47.7	226	7	US-11-096-568A-18946	Sequence 18946, A	362	31	47.7	736	7	US-11-085-185-4	Sequence 4, Appl
290	31	47.7	227	7	US-11-170-653-21	Sequence 21, Appl	363	31	47.7	783	7	US-11-082-389-354	Sequence 354, App
291	31	47.7	227	7	US-11-170-653-27	Sequence 27, Appl	364	31	47.7	867	6	US-10-467-657-2208	Sequence 2208, Ap
292	31	47.7	231	7	US-11-096-568A-23503	Sequence 23503, A	365	31	47.7	937	7	US-11-098-686-11296	Sequence 11296, A
293	31	47.7	239	7	US-11-096-568A-4422	Sequence 4422, Ap	366	31	47.7	1142	6	US-10-501-035-376	Sequence 376, App
294	31	47.7	243	7	US-11-096-568A-4421	Sequence 4421, Ap	367	31	47.7	1142	6	US-11-109-156-22	Sequence 22, Appl
295	31	47.7	246	7	US-11-134-703-10	Sequence 10, Appl	368	31	47.7	1154	6	US-10-330-773-867	Sequence 867, App
296	31	47.7	246	7	US-11-134-703-12	Sequence 12, Appl	369	31	47.7	1170	6	US-10-330-773-864	Sequence 864, App
297	31	47.7	258	7	US-11-098-686-10488	Sequence 10488, A	370	31	47.7	2204	6	US-10-495-083-8	Sequence 8, Appl
298	31	47.7	265	7	US-11-096-568A-16347	Sequence 16347, A	371	31	47.7	2479	7	US-11-087-099-2391	Sequence 2391, Ap
299	31	47.7	265	7	US-11-096-568A-29462	Sequence 29462, A	372	31	47.7	7968	7	US-11-143-980-49	Sequence 49, Appl
300	31	47.7	266	7	US-11-096-568A-33429	Sequence 33429, A	373	30.5	46.9	119	59	US-11-096-568A-13890	Sequence 13890, A
301	31	47.7	267	7	US-11-096-568A-29461	Sequence 29461, A	374	30.5	46.9	129	7	US-11-096-568A-13889	Sequence 13889, A
302	31	47.7	268	7	US-11-134-703-4	Sequence 4, Appl	375	30.5	46.9	204	7	US-11-116-144-149	Sequence 149, App
303	31	47.7	268	7	US-11-134-703-18	Sequence 18, Appl	376	30.5	46.9	204	7	US-11-062-471A-1	Sequence 1, Appl
304	31	47.7	277	7	US-11-140-416-34	Sequence 34, Appl	377	30.5	46.9	229	7	US-11-062-471A-2	Sequence 2, Appl
305	31	47.7	277	7	US-11-096-568A-29460	Sequence 29460, A	378	30.5	46.9	229	7	US-11-062-471A-9	Sequence 9, Appl
306	31	47.7	289	7	US-11-096-568A-16829	Sequence 16829, A	379	30.5	46.9	264	6	US-11-087-099-7999	Sequence 7999, Ap
307	31	47.7	289	7	US-11-096-568A-18354	Sequence 18354, A	380	30.5	46.9	264	6	US-10-995-561-740	Sequence 740, App
308	31	47.7	297	7	US-11-096-568A-18353	Sequence 18353, A	381	30.5	46.9	344	6	US-10-995-561-744	Sequence 744, App
309	31	47.7	297	7	US-11-096-568A-25291	Sequence 25291, A	382	30.5	46.9	391	6	US-10-995-561-739	Sequence 739, App
310	31	47.7	298	7	US-11-096-568A-25290	Sequence 25290, A	383	30.5	46.9	432	6	US-10-995-561-738	Sequence 738, App
311	31	47.7	299	7	US-11-096-568A-25289	Sequence 25289, A	384	30.5	46.9	457	6	US-10-995-561-741	Sequence 741, App
312	31	47.7	301	6	US-10-331-826A-176	Sequence 176, App	385	30.5	46.9	491	6	US-10-995-561-743	Sequence 743, App
313	31	47.7	301	6	US-10-973-115B-176	Sequence 176, App	386	30.5	46.9	512	6	US-10-995-561-745	Sequence 745, App
314	31	47.7	308	7	US-11-096-568A-15851	Sequence 15851, A	387	30.5	46.9	601	6	US-10-995-561-743	Sequence 903, App
315	31	47.7	316	7	US-11-024-959-453	Sequence 453, App	388	30.5	46.9	609	7	US-11-087-099-903	Sequence 903, App
316	31	47.7	319	7	US-11-096-568A-18352	Sequence 18352, A	389	30.5	46.9	1059	7	US-11-087-099-7286	Sequence 7286, Ap
317	31	47.7	321	7	US-11-024-959-333	Sequence 333, App	390	30.5	46.9	1067	7	US/11/062	Sequence 3, Appl

391	30.5	46.9	1070	7	US-11-062	Sequence 4, Appli	464	30	46.2	237	7	US-11-194-246-302	Sequence 302, App
392	30.5	46.9	1084	7	US-11-062	Sequence 8, Appli	465	30	46.2	243	7	US-11-087-099-3396	Sequence 3396, Ap
393	30.5	46.9	1092	7	US-11-062	Sequence 6, Appli	466	30	46.2	245	7	US-11-054-515-1864	Sequence 1864, Ap
394	30.5	46.9	1095	7	US-11-062	Sequence 7, Appli	467	30	46.2	250	6	US-10-131-826A-78	Sequence 78, Appl
395	30	46.2	12	7	US-11-098-763-15	Sequence 15, Appl	468	30	46.2	250	6	US-10-973-115B-78	Sequence 78, Appl
396	30	46.2	15	6	US-10-718-264-59	Sequence 59, Appl	469	30	46.2	250	7	US-11-096-568A-20867	Sequence 20867, A
397	30	46.2	15	6	US-10-718-264-149	Sequence 149, App	470	30	46.2	251	7	US-11-096-568A-26008	Sequence 26008, A
398	30	46.2	15	6	US-10-718-264-150	Sequence 150, App	471	30	46.2	252	7	US-11-087-099-392	Sequence 392, App
399	30	46.2	15	6	US-10-718-264-59	Sequence 59, Appl	472	30	46.2	252	7	US-11-096-568A-32817	Sequence 32817, A
400	30	46.2	15	6	US-10-718-264-149	Sequence 149, App	473	30	46.2	254	7	US-11-087-099-8546	Sequence 8546, Ap
401	30	46.2	15	6	US-10-718-264-150	Sequence 150, App	474	30	46.2	258	7	US-11-096-568A-30003	Sequence 30003, A
402	30	46.2	16	7	US-11-009-939-34	Sequence 34, Appl	475	30	46.2	260	7	US-11-096-568A-9211	Sequence 9211, Ap
403	30	46.2	16	6	US-10-895-064-143	Sequence 143, App	476	30	46.2	262	7	US-11-096-568A-3056	Sequence 3056, Ap
404	30	46.2	18	7	US-11-128-741-143	Sequence 143, App	477	30	46.2	262	7	US-11-096-568A-3057	Sequence 3057, Ap
405	30	46.2	18	7	US-11-128-741-143	Sequence 143, App	478	30	46.2	263	6	US-10-821-234-1571	Sequence 1571, Ap
406	30	46.2	18	7	US-11-128-741-143	Sequence 143, App	479	30	46.2	263	6	US-10-878-568A-128	Sequence 128, App
407	30	46.2	18	7	US-11-235-009-12	Sequence 12, Appl	480	30	46.2	263	7	US-11-096-568A-9210	Sequence 9210, Ap
408	30	46.2	90	7	US-11-082-389-360	Sequence 360, App	481	30	46.2	264	7	US-11-096-568A-30608	Sequence 30608, A
409	30	46.2	98	7	US-11-144-248-34	Sequence 34, Appl	482	30	46.2	266	7	US-11-096-568A-3055	Sequence 3055, Ap
410	30	46.2	98	7	US-11-054-669-34	Sequence 34, Appl	483	30	46.2	268	7	US-11-096-568A-3055	Sequence 3055, Ap
411	30	46.2	98	7	US-11-144-222-34	Sequence 34, Appl	484	30	46.2	271	7	US-11-058-817A-4	Sequence 4, Appli
412	30	46.2	98	7	US-11-004-590-39	Sequence 39, Appl	485	30	46.2	271	7	US-11-096-568A-19845	Sequence 19845, A
413	30	46.2	98	7	US-11-182-343-34	Sequence 34, Appl	486	30	46.2	272	7	US-11-058-817A-6	Sequence 6, Appli
414	30	46.2	109	7	US-11-072-512-3806	Sequence 3806, Ap	487	30	46.2	273	7	US-11-096-568A-21721	Sequence 21721, A
415	30	46.2	117	7	US-11-049-536-634	Sequence 634, App	488	30	46.2	276	7	US-11-087-099-10298	Sequence 10298, A
416	30	46.2	117	7	US-11-096-568A-11	Sequence 11, Appl	489	30	46.2	277	7	US-11-096-568A-21095	Sequence 21095, A
417	30	46.2	118	7	US-11-049-536-334	Sequence 334, App	490	30	46.2	278	7	US-11-096-568A-32957	Sequence 32957, A
418	30	46.2	118	7	US-11-049-536-334	Sequence 334, App	491	30	46.2	280	7	US-11-096-568A-9209	Sequence 9209, Ap
419	30	46.2	119	7	US-11-049-536-534	Sequence 534, App	492	30	46.2	282	6	US-10-467-657-4418	Sequence 4418, Ap
420	30	46.2	120	7	US-11-049-536-650	Sequence 650, App	493	30	46.2	282	7	US-11-096-568A-23339	Sequence 23339, A
421	30	46.2	120	7	US-11-096-568A-30949	Sequence 30949, A	494	30	46.2	284	7	US-11-096-568A-11255	Sequence 11255, A
422	30	46.2	121	7	US-11-009-939-32	Sequence 32, Appl	495	30	46.2	285	7	US-11-096-568A-33279	Sequence 33279, A
423	30	46.2	123	6	US-11-087-099-8359	Sequence 8359, Ap	496	30	46.2	288	6	US-10-467-657-4590	Sequence 4590, Ap
424	30	46.2	124	6	US-10-925-366A-211	Sequence 211, App	497	30	46.2	290	7	US-11-087-099-9971	Sequence 9971, Ap
425	30	46.2	124	7	US-11-096-568A-14160	Sequence 14160, A	498	30	46.2	293	7	US-11-087-099-9115	Sequence 9115, Ap
426	30	46.2	125	7	US-11-049-536-478	Sequence 478, App	499	30	46.2	300	7	US-11-096-568A-5688	Sequence 5688, Ap
427	30	46.2	130	6	US-10-993-543-144	Sequence 144, App	500	30	46.2	302	7	US-11-087-099-3041	Sequence 3041, Ap
428	30	46.2	133	5	US-09-978-360A-415	Sequence 415, App							
429	30	46.2	136	7	US-11-087-099-700	Sequence 700, App							
430	30	46.2	136	7	US-11-096-568A-25660	Sequence 25660, A							
431	30	46.2	142	7	US-11-096-568A-10	Sequence 10, Appl							
432	30	46.2	147	7	US-11-087-099-9798	Sequence 9798, Ap							
433	30	46.2	152	7	US-11-087-099-8123	Sequence 8123, Ap							
434	30	46.2	152	7	US-11-096-568A-7556	Sequence 7556, Ap							
435	30	46.2	152	7	US-11-096-568A-30948	Sequence 30948, A							
436	30	46.2	153	7	US-11-150-533-16	Sequence 16, Appl							
437	30	46.2	163	6	US-11-102-240-160	Sequence 160, App							
438	30	46.2	166	7	US-11-096-568A-9	Sequence 9, Appli							
439	30	46.2	167	7	US-11-096-568A-23341	Sequence 23341, A							
440	30	46.2	191	7	US-11-087-099-1813	Sequence 1813, Ap							
441	30	46.2	191	7	US-11-087-099-8458	Sequence 8458, Ap							
442	30	46.2	191	7	US-11-096-568A-20264	Sequence 20264, A							
443	30	46.2	192	7	US-11-096-568A-30610	Sequence 30610, A							
444	30	46.2	194	7	US-11-087-099-10291	Sequence 10291, A							
445	30	46.2	197	7	US-11-096-568A-24258	Sequence 24258, A							
446	30	46.2	201	7	US-11-096-568A-21722	Sequence 21722, A							
447	30	46.2	202	7	US-11-087-099-10401	Sequence 10401, A							
448	30	46.2	208	7	US-11-096-568A-19846	Sequence 19846, A							
449	30	46.2	211	7	US-11-087-099-1427	Sequence 1427, Ap							
450	30	46.2	214	7	US-11-096-568A-30004	Sequence 30004, A							
451	30	46.2	214	7	US-11-087-099-1394	Sequence 1394, Ap							
452	30	46.2	217	7	US-11-087-099-8188	Sequence 8188, Ap							
453	30	46.2	217	7	US-11-096-568A-23340	Sequence 23340, A							
454	30	46.2	219	7	US-11-194-246-339	Sequence 339, App							
455	30	46.2	223	7	US-11-096-568A-3514	Sequence 3514, Ap							
456	30	46.2	225	6	US-10-517-939-172	Sequence 172, App							
457	30	46.2	228	7	US-11-096-568A-3513	Sequence 3513, Ap							
458	30	46.2	229	7	US-11-096-568A-30153	Sequence 30153, A							
459	30	46.2	231	7	US-11-087-099-9419	Sequence 9419, Ap							
460	30	46.2	232	7	US-11-055-822-848	Sequence 848, App							
461	30	46.2	233	6	US-10-718-264-12	Sequence 12, Appl							
462	30	46.2	233	6	US-10-718-264-12	Sequence 12, Appl							
463	30	46.2	233	6	US-10-718-264-12	Sequence 12, Appl							

ALIGNMENTS

RESULT 1
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760

```
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match      100.0%; Score 65; DB 7; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.005; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY      1 YALWSPNGK 10
        |||||
Db      210 YALWSPNGK 219

RESULT 2
US-11-116-939-13
; Sequence 13, Application US/11116919
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.0115U2
; CURRENT APPLICATION NUMBER: US/11/116,939
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; PRIOR FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
US-11-116-939-13

Query Match      81.5%; Score 53; DB 7; Length 762;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
        |||||
Db      209 ALWWSPPNG 216

RESULT 3
US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Query Match      81.5%; Score 53; DB 6; Length 766;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
        |||||
Db      213 ALWWSPPNG 220
```

```
RESULT 4
US-10-501-035-234
; Sequence 234, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASES
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-234

Query Match      81.5%; Score 53; DB 6; Length 766;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 9
        |||||
Db      213 ALWWSPPNG 220

RESULT 5
US-11-090-439-58
; Sequence 58, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel P.
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-58

Query Match      66.2%; Score 43; DB 7; Length 582;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ALWWSPPNG 10
        |||||
Db      418 ALFWMPNGK 426

RESULT 6
US-11-169-041-130
; Sequence 130, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
```

```
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-169-041-130

Query Match 66.2%; Score 43; DB 7; Length 582;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSPNGK 10
Db 418 ALFWMPNGK 426

RESULT 7
US-11-200-822-2
; Sequence 2, Application US/11200822
; Publication No. US20060040302A1
; GENERAL INFORMATION:
; APPLICANT: Bostein, et al.
; TITLE OF INVENTION: Methods of Classifying, Diagnosing, Stratifying and
; TITL OF INVENTION: Treating Cancer Patients and Their Tumors
; FILE REFERENCE: 2002850-0049
; CURRENT APPLICATION NUMBER: US/11/200,822
; CURRENT FILING DATE: 2005-08-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Matrix
US-11-200-822-2

Query Match 66.2%; Score 43; DB 7; Length 582;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSPNGK 10
Db 418 ALFWMPNGK 426

RESULT 8
US-11-024-959-477
; Sequence 477, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30

; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 477
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-477

Query Match 63.1%; Score 41; DB 7; Length 964;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWWSPNGK 10
Db 244 SLWSPNGK 252

RESULT 9
US-11-087-099-7087
; Sequence 7087, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7087
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Oenococcus oeni
US-11-087-099-7087

Query Match 61.5%; Score 40; DB 7; Length 475;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YALWWSPNGK 10
Db 193 HGFWYTPNGK 192

RESULT 10
US-10-821-234-1610
; Sequence 1610, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Suean
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1610
; LENGTH: 2335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1610

Query Match 61.5%; Score 40; DB 6; Length 2335;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YALWWSNP 7
Db 1534 FTLWWSNP 1540
```

RESULT 11
US-11-009-658-16
; Sequence 16, Application US/11009658
; Publication No. US20060003430A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Lysophosphatidic Acid Acetyltransferases
; FILE REFERENCE: BB1332
; CURRENT APPLICATION NUMBER: US/11/009,658
; PRIOR FILING DATE: 2004-12-10
; PRIOR APPLICATION NUMBER: US/09/914,098
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/121,119
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (136)
; OTHER INFORMATION: Xaa = any amino acid
US-11-009-658-16

Query Match 60.0%; Score 39; DB 7; Length 174;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WSPNG 9
|||||
DB 67 WSPRG 72

RESULT 12
US-11-009-658-42
; Sequence 42, Application US/11009658
; Publication No. US20060003430A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Lysophosphatidic Acid Acetyltransferases
; FILE REFERENCE: BB1332
; CURRENT APPLICATION NUMBER: US/11/009,658
; CURRENT FILING DATE: 2004-12-10
; PRIOR APPLICATION NUMBER: US/09/914,098
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/121,119
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 42
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-009-658-42

Query Match 60.0%; Score 39; DB 7; Length 255;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WSPNG 9
|||||
DB 67 WSPRG 72

RESULT 13
US-11-024-959-470
; Sequence 470, Application US/11024959
; Publication No. US2006010516A1
; GENERAL INFORMATION:

; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 470
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-470

Query Match 60.0%; Score 39; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WSPNG 10
|||||
DB 88 WSPNG 93

RESULT 14
US-10-995-561-870
; Sequence 870, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 870
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-870

Query Match 60.0%; Score 39; DB 6; Length 1234;
Best Local Similarity 87.5%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALWSPNG 9
|||||
DB 132 ALWSPNG 139

RESULT 15
US-10-995-561-867
; Sequence 867, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 867
; LENGTH: 1365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-867

Query Match 60.0%; Score 39; DB 6; Length 1365;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
|||
Db 132 ALWSPNG 139

RESULT 16
US-10-995-561-868
; Sequence 868, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 868
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-868

Query Match 60.0%; Score 39; DB 6; Length 1366;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
|||
Db 132 ALWSPNG 139

RESULT 17
US-10-995-561-869
; Sequence 869, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 869
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-869

Query Match 60.0%; Score 39; DB 6; Length 1411;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ALWSPNG 9
|||
Db 132 ALWSPNG 139

RESULT 18
US-11-072-512-2562
; Sequence 2562, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2562
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2562

Query Match 58.5%; Score 38; DB 7; Length 104;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LWSPNGK 10
|||
Db 51 LWDPYQG 58

RESULT 19
US-11-098-686-10864
; Sequence 10864, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10864
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10864

Query Match 58.5%; Score 38; DB 7; Length 228;

Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LWSPNG 8
|||
Db 54 LWSPD 59

RESULT 20
US-11-116-144-91
; Sequence 91, Application US/11116144
; Publication No. US20050277181A1
; GENERAL INFORMATION:
; APPLICANT: BERTHET, FRANCOIS XAVIER
; APPLICANT: CASADEVALL, FRANCESC VAYREDA
; APPLICANT: SANZ MARIA, MARIA CRUZ
; APPLICANT: GARCIA, TERESA LLOP
; APPLICANT: OLLE, ANGELS MOR
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING PATHOGEN
; FILE REFERENCE: INL-084
; CURRENT APPLICATION NUMBER: US/11/116,144
; PRIOR FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: PCT/ES04/000581
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: EP 03380307.3
; PRIOR FILING DATE: 2003-12-23
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 91
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Anas platyrhynchos
US-11-116-144-91

Query Match 56.9%; Score 37; DB 7; Length 127;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 WSPNGK 10
|||
Db 60 WWCNGK 66

RESULT 21
US-11-116-144-90
; Sequence 90, Application US/11116144
; Publication No. US20050277181A1
; GENERAL INFORMATION:
; APPLICANT: BERTHET, FRANCOIS XAVIER
; APPLICANT: CASADEVALL, FRANCESC VAYREDA
; APPLICANT: SANZ MARIA, MARIA CRUZ
; APPLICANT: GARCIA, TERESA LLOP
; APPLICANT: OLLE, ANGELS MOR
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING PATHOGEN
; FILE REFERENCE: INL-084
; CURRENT APPLICATION NUMBER: US/11/116,144
; PRIOR FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: PCT/ES04/000581
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: EP 03380307.3
; PRIOR FILING DATE: 2003-12-23
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 90
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Anas platyrhynchos
US-11-116-144-90

Query Match 56.9%; Score 37; DB 7; Length 129;
Best Local Similarity 71.4%; Pred. No. 27;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 WSPNGK 10
|||
Db 62 WWCNGK 68

RESULT 22
US-11-096-568A-21612
; Sequence 21612, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21612
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(261)
; OTHER INFORMATION: Ceres Seq. ID no. 12405497
US-11-096-568A-21612

Query Match 56.9%; Score 37; DB 7; Length 261;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWSPNG 9
|||
Db 138 ALWWSAYG 145

RESULT 23
US-11-096-568A-21611
; Sequence 21611, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21611
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(271)
; OTHER INFORMATION: Ceres Seq. ID no. 12405496
US-11-096-568A-21611

Query Match 56.9%; Score 37; DB 7; Length 271;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWSPNG 9
|||
Db 148 ALWWSAYG 155

RESULT 24
US-11-096-568A-10627
; Sequence 10627, Application US/11096568A
; Publication No. US20060048240A1

```
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 10627
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(272)
; OTHER INFORMATION: Ceres Seq. ID no. 13596638
US-11-096-568A-10627

Query Match          56.9%; Score 37; DB 7; Length 272;
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 ALWSPNG 9
      |||||
Db      148 ALWSSYG 155

RESULT 25
US-11-096-568A-21610
; Sequence 21610, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21610
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(327)
; OTHER INFORMATION: Ceres Seq. ID no. 12405495
US-11-096-568A-21610

Query Match          56.9%; Score 37; DB 7; Length 327;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 ALWSPNG 9
      |||||
Db      204 ALWBSYG 211

RESULT 26
US-11-096-568A-10626
; Sequence 10626, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 10626
; LENGTH: 328
; TYPE: PRT
```

```
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(328)
; OTHER INFORMATION: Ceres Seq. ID no. 13596637
US-11-096-568A-10626

Query Match          56.9%; Score 37; DB 7; Length 328;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 ALWSPNG 9
      |||||
Db      204 ALWSSYG 211

RESULT 27
US-11-096-568A-10625
; Sequence 10625, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 10625
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(370)
; OTHER INFORMATION: Ceres Seq. ID no. 13596636
US-11-096-568A-10625

Query Match          56.9%; Score 37; DB 7; Length 370;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 ALWSPNG 9
      |||||
Db      246 ALWSSYG 253

RESULT 28
US-11-096-568A-17843
; Sequence 17843, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 17843
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(606)
; OTHER INFORMATION: Ceres Seq. ID no. 12361175
US-11-096-568A-17843

Query Match          56.9%; Score 37; DB 7; Length 606;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 3 LWMSPNG 9
|||
Db 181 LWMFPG 187

RESULT 29
US-11-024-959-381
; Sequence 381, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 381
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-381

Query Match 56.9%; Score 37; DB 7; Length 607;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
|||
Db 239 YAVNSPDSK 248

RESULT 30
US-11-107-096-51
; Sequence 51, Application US/11107096
; Publication No. US20060003348A1
; GENERAL INFORMATION:
; APPLICANT: SIDHU, SACHDEV S.
; APPLICANT: ZHANG, YINGNAN
; TITLE OF INVENTION: OMI PDZ MODULATORS
; FILE REFERENCE: P2100R1
; CURRENT APPLICATION NUMBER: US/11/107,096
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 60/563,157
; PRIOR FILING DATE: 2004-04-16
; NUMBER OF SEQ ID NOS: 79
; SEQ ID NO 51
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: sequence is synthesized
US-11-107-096-51

Query Match 55.4%; Score 36; DB 7; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
:
Db 1 WGYWMPGPG 9

RESULT 31

US-11-096-568A-3423
; Sequence 3423, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 3423
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(209)
; OTHER INFORMATION: Ceres Seq. ID no. 12608263
US-11-096-568A-3423

Query Match 55.4%; Score 36; DB 7; Length 209;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WMSPNGK 10
:
Db 17 FWKPNGK 23

RESULT 32
US-10-467-657-5656
; Sequence 5656, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: Seqwin99, version 1.04
; SEQ ID NO 5656
; LENGTH: 929
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5656

Query Match 55.4%; Score 36; DB 6; Length 929;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
:
Db 672 YGKWMDLNG 680

RESULT 33
US-11-049-536-84
; Sequence 84, Application US/11049536
; Publication No. US20060024297A1
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive R.
; APPLICANT: Dransfield, Daniel T.
; APPLICANT: Pieters, Henk
; APPLICANT: Hoet, Rene
; APPLICANT: Hufton, Simon E.

```
; TITLE OF INVENTION: TIE COMPLEX BINDING PROTEINS
; FILE REFERENCE: 10280-128001
; CURRENT APPLICATION NUMBER: US/11/049,536
; CURRENT FILING DATE: 2005-02-02
; PRIOR APPLICATION NUMBER: US 10/916,840
; PRIOR FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/494,713
; PRIOR FILING DATE: 2003-08-12
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-11-049-536-84

Query Match      54.6%; Score 35.5; DB 7; Length 131;
Best Local Similarity 29.2%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 0; Indels 15; Gaps 1;

Qy      1 YALMW-----SPNG 9
Db      32 YAMWVRQAPGKGLEWVSSISPSG 55
      |||:|
      |||:|

RESULT 34
US-11-112-882-32
; Sequence 32, Application US/11112882
; Publication No. US20050273885A1
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific and Industrial Research Organisation
; TITLE OF INVENTION: Synthesis of Long-Chain Polyunsaturated Fatty Acids in Recombinant
; FILE REFERENCE: 503244
; CURRENT APPLICATION NUMBER: US/11/112,882
; CURRENT FILING DATE: 2005-04-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-11-112-882-32

Query Match      54.6%; Score 35.5; DB 7; Length 318;
Best Local Similarity 33.3%; Pred. No. 1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 5; Gaps 1;

Qy      1 YALMW-----SPNGK 10
Db      192 FTIMLVTFVAPNGE 206
      : :|
      : :|

RESULT 35
US-10-925-366A-352
; Sequence 352, Application US/10925366A
; Publication No. US20050271663A1
; GENERAL INFORMATION:
; APPLICANT: Ignatovich, Olga
; APPLICANT: Dewiltdt, Rudolph M.T.
; APPLICANT: Benjamin, Woolven
; APPLICANT: Grant, Steven
; APPLICANT: Jones, Philip
; APPLICANT: Bastran, Amrik
; APPLICANT: Brewis, Neil
; TITLE OF INVENTION: Compositions and Methods for Treating Inflammatory Disorders
; FILE REFERENCE: 8039/2105
; CURRENT APPLICATION NUMBER: US/10/925,366A
; CURRENT FILING DATE: 2004-08-24
; PRIOR APPLICATION NUMBER: US 10/744,774
; PRIOR FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: PCT/GB2003/002804
```

```
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: PCT/GB2002/03014
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: GB 0230202.4
; PRIOR FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: GB 115841.9
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PCT/GB2004/002829
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/535,076
; PRIOR FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: PCT/GB2003/005646
; PRIOR FILING DATE: 2003-12-24
; PRIOR APPLICATION NUMBER: GB 0327706.8
; PRIOR FILING DATE: 2003-11-28
; PRIOR APPLICATION NUMBER: US 60/509,613
; PRIOR FILING DATE: 2003-10-08
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 352
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino Acid Sequence of Domain Antibody
US-10-925-366A-352

Query Match      53.8%; Score 35; DB 6; Length 119;
Best Local Similarity 80.0%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YALMW 5
Db      32 YAMWM 36
      |||:|
      |||:|

RESULT 36
US-10-993-543-180
; Sequence 180, Application US/10993543
; Publication No. US20060036076A1
; GENERAL INFORMATION:
; APPLICANT: Dransfield, Daniel T.
; APPLICANT: Rookey, Kristin
; APPLICANT: Ladner, Robert C.
; TITLE OF INVENTION: Metalloproteinase-Binding Proteins
; FILE REFERENCE: 10280-088001
; CURRENT APPLICATION NUMBER: US/10/993,543
; CURRENT FILING DATE: 2004-11-19
; PRIOR APPLICATION NUMBER: US 60/523,745
; PRIOR FILING DATE: 2003-11-19
; NUMBER OF SEQ ID NOS: 299
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 180
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-993-543-180

Query Match      53.8%; Score 35; DB 6; Length 129;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YALMW 5
Db      32 YAMWM 36
      |||:|
      |||:|

RESULT 37
US-11-098-686-10681
; Sequence 10681, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
```

; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; CURRENT APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10681
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10681

Query Match 53.8%; Score 35; DB 7; Length 282;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 YALWSPNG 9
: ||| :|||
Db 195 FTLWETPENG 203

RESULT 38
US-10-063-703-166
; Sequence 166, Application US/10063703
; Publication No. US20060008901A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,703
; PRIOR FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 166
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-703-166

Query Match 53.8%; Score 35; DB 6; Length 303;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ALWWSVP 7
: ||| :|||
Db 16 ALWWSVP 21

RESULT 39
US-11-102-240-166
; Sequence 166, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS

; TITLE OF INVENTION: ESOPHAGEAL TUMOR
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 166
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-166

Query Match 53.8%; Score 35; DB 7; Length 303;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ALWWSVP 7
: ||| :|||
Db 16 ALWWSVP 21

RESULT 40
US-11-241-056-12
; Sequence 12, Application US/11241056
; Publication No. US20060024807A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-US
; CURRENT APPLICATION NUMBER: US/11/241,056
; CURRENT FILING DATE: 2005-09-30
; PRIOR APPLICATION NUMBER: US/09/980,464
; PRIOR FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-241-056-12

Query Match 53.8%; Score 35; DB 7; Length 311;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 YALWWSVP 7
: ||| :|||
Db 268 HVLWVKP 274

RESULT 41
US-11-098-686-10683
; Sequence 10683, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318

; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10683
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10683

Query Match 53.8%; Score 35; DB 7; Length 441;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
: ||| :
Db 353 FTLWETPG 361

RESULT 42
US-10-517-939-368
; Sequence 368, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza

; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; PRIOR FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 368
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(23)
US-10-517-939-368

Query Match 53.8%; Score 35; DB 6; Length 445;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 YALWSPNG 9
: ||| :
Db 38 YTWSDGG 46

RESULT 43
US-11-024-959-507
; Sequence 507, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.

; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 507
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-507

Query Match 53.8%; Score 35; DB 7; Length 463;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALWSPNGK 10
: ||| :
Db 237 ALSWPTGR 245

RESULT 44
US-10-517-939-250
; Sequence 250, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
; APPLICANT: Steer, Brian
; APPLICANT: Callen, Walter
; APPLICANT: Healey, Shaun
; APPLICANT: Hazlewood, Geoff
; APPLICANT: Wu, Di
; APPLICANT: Blum, David
; APPLICANT: Esteghlalian, Alireza
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 564462007901
; CURRENT APPLICATION NUMBER: US/10/517,939
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: PCT/US03/19153
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; NAME/KEY: SIGNAL
; LOCATION: (1)...(33)
US-10-517-939-250

Query Match 53.8%; Score 35; DB 6; Length 479;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 YALWSPNGK 10
: ||| :
Db 98 YSDWADPGK 107

RESULT 45
US-11-072-512-3591
; Sequence 3591, Application US/11072512
; Publication No. US20060029945A1

```

; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIRO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3591
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-072-512-3591

```

```

Query Match. 53.8%; Score 35; DB 7; Length 484;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 LWMSPNGK 10
   ||:||||
DB 169 IWRTPNGK 176

```

```

RESULT 46
US-11-087-099-9085
; Sequence 9085, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9085
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Nicotiana plumbaginifolia
; US-11-087-099-9085

```

```

Query Match 53.8%; Score 35; DB 7; Length 508;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 YALWSP 7
   ||:||||
DB 24 YTIWWP 30

```

```

RESULT 47
US-11-087-099-10512
; Sequence 10512, Application US/11087099

```

```

; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10512
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Neurospora crassa
; US-11-087-099-10512

```

```

Query Match 53.8%; Score 35; DB 7; Length 547;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 LWMSPNG 9
   ||:||||
DB 88 IWMTVNG 94

```

```

RESULT 48
US-11-080-991-88
; Sequence 88, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-080-991-88

```

```

Query Match 53.8%; Score 35; DB 7; Length 607;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 ALWSPNGK 10
   ||:||||
DB 442 AIWWDVGK 450

```

```

RESULT 49
US-11-024-959-481
; Sequence 481, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30

```


; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 481
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-481

Query Match 53.8%; Score 35; DB 7; Length 608;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VALWSPNGK 10
||| :||| :
Db 237 YAAWSWSPDSK 246

RESULT 50

US-11-072-512-3453

; Sequence 3453, Application US/11072512
; Publication No. US20060029945A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHIKO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOTYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: Novel full length cDNA

; FILE REFERENCE: 084335-0191

; CURRENT APPLICATION NUMBER: US/11/072,512

; CURRENT FILING DATE: 2005-03-07

; PRIOR APPLICATION NUMBER: US 60/350,978

; PRIOR FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: JP 2001-379298

; PRIOR FILING DATE: 2001-11-05

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3453

; LENGTH: 791

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-072-512-3453

Query Match 53.8%; Score 35; DB 7; Length 791;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LWWSPNGK 10
|| :||| :
Db 476 IWRTPNGK 483

RESULT 51

US-10-613-744-41

; Sequence 41, Application US/10613744

; Publication No. US20050272093A1

; GENERAL INFORMATION:

; APPLICANT: MacKinnon, Roderick

; APPLICANT: The Rockefeller University

; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With

; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 41

; LENGTH: 36

; TYPE: PRT

; ORGANISM: Streptomyces lividans

US-10-613-744-41

Query Match 52.3%; Score 34; DB 6; Length 36;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWS 6

|||||

Db 15 ALWWS 19

|||||

RESULT 52

US-10-613-744-23

; Sequence 23, Application US/10613744

; Publication No. US20050272093A1

; GENERAL INFORMATION:

; APPLICANT: MacKinnon, Roderick

; APPLICANT: The Rockefeller University

; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With

; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation

; TITLE OF INVENTION: Channel Proteins, and Uses Thereof

; FILE REFERENCE: 018512-002901US

; CURRENT APPLICATION NUMBER: US/10/613,744

; CURRENT FILING DATE: 2003-07-03

; PRIOR APPLICATION NUMBER: US/09/275,252

; PRIOR FILING DATE: 1999-03-24

; PRIOR APPLICATION NUMBER: US 09/045,529

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: US 09/054,347

; PRIOR FILING DATE: 1998-04-02

; PRIOR APPLICATION NUMBER: WO PCT/US99/06307

; PRIOR FILING DATE: 1999-03-22

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 23

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Streptomyces lividans

US-10-613-744-23

Query Match 52.3%; Score 34; DB 6; Length 40;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALWWS 6

|||||

Db 16 ALWWS 20

|||||

RESULT 53

US-10-613-744-3

; Sequence 3, Application US/10613744

; Publication No. US20050272093A1

; GENERAL INFORMATION:

; APPLICANT: MacKinnon, Roderick

; APPLICANT: The Rockefeller University

; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With

Mon Mar 13 10:54:36 2006

APPLICANT: MacKinnon, Roderick
APPLICANT: The Rockefeller University
TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
TITLE OF INVENTION: Channel Proteins, and Uses Thereof
FILE REFERENCE: 018512-002901US
CURRENT APPLICATION NUMBER: US/10/613,744
CURRENT FILING DATE: 2003-07-03
PRIOR APPLICATION NUMBER: US/09/275,252
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: US 09/045,529
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: US 09/054,347
PRIOR FILING DATE: 1998-04-02
PRIOR APPLICATION NUMBER: WO PCT/US99/06307
PRIOR FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
TYPE: PRT
ORGANISM: Clostridium acetobutylicum
US-10-613-744-3

Query Match 52.3%; Score 34; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWS 6
DB 5 ALWWS 9

RESULT 54
US-10-613-744-25
Sequence 25, Application US/10613744
Publication No. US20050272093A1
GENERAL INFORMATION:
APPLICANT: MacKinnon, Roderick
APPLICANT: The Rockefeller University
TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
TITLE OF INVENTION: Channel Proteins, and Uses Thereof
FILE REFERENCE: 018512-002901US
CURRENT APPLICATION NUMBER: US/10/613,744
CURRENT FILING DATE: 2003-07-03
PRIOR APPLICATION NUMBER: US/09/275,252
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: US 09/045,529
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: US 09/054,347
PRIOR FILING DATE: 1998-04-02
PRIOR APPLICATION NUMBER: WO PCT/US99/06307
PRIOR FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 25
LENGTH: 58
TYPE: PRT
ORGANISM: Streptomyces lividans
US-10-613-744-25

Query Match 52.3%; Score 34; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALWWS 6
DB 5 ALWWS 9

RESULT 55
US-11-096-568A-27334

Sequence 27334, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Therby
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 27334
LENGTH: 91
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(91)
OTHER INFORMATION: Ceres Seq. ID no. 13648513
US-11-096-568A-27334

Query Match 52.3%; Score 34; DB 7; Length 91;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWSNP 8
DB 69 WWSNP 73

RESULT 56
US-11-087-099-4615
Sequence 4615, Application US/11087099
Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)B EP
CURRENT APPLICATION NUMBER: US/11/087,099
CURRENT FILING DATE: 2005-03-22
NUMBER OF SEQ ID NOS: 12464
SEQ ID NO 4615
LENGTH: 134
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(134)
OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-4615

Query Match 52.3%; Score 34; DB 7; Length 134;
Best Local Similarity 83.3%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LWSNP 8
DB 107 LWSNP 112

RESULT 57
US-11-096-568A-15758
Sequence 15758, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Therby
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 15758
LENGTH: 156

```
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(156)
; OTHER INFORMATION: Ceres Seq. ID no. 12348098
US-11-096-568A-15758

Query Match      52.3%; Score 34; DB 7; Length 156;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 WMSPNG 9
Db      115 WWAPAG 120

RESULT 58
US-10-613-744-1
; Sequence 1, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact with
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US/09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US/09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Streptomyces lividans
US-10-613-744-1

Query Match      52.3%; Score 34; DB 6; Length 160;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ALWWS 6
Db      65 ALWWS 69

RESULT 59
US-10-613-744-16
; Sequence 16, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact with
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US/09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US/09/054,347
```

```
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 16
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mutated
; OTHER INFORMATION: prokaryotic kcsa potassium channel protein of
; OTHER INFORMATION: Streptomyces lividans
US-10-613-744-16

Query Match      52.3%; Score 34; DB 6; Length 160;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ALWWS 6
Db      65 ALWWS 69

RESULT 60
US-11-072-512-2677
; Sequence 2677, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2677
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2677

Query Match      52.3%; Score 34; DB 7; Length 181;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 WMSPNG 9
Db      25 WWAPMG 30

Search completed: March 11, 2006, 12:32:58
```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:11:52 ; Search time 140.132 Seconds
(without alignments)
47.032 Million cell updates/sec

Title: US-10-774-242A-1

Perfect score: 76

Sequence: 1 IVLPRSRVHNEENT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 500 summaries

Database :

A_Geneseq_21.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	15	8	Adi47257 Human alp
2	76	100.0	50	9	Aeb94176 Human sol
3	76	100.0	759	2	Aaw31963 Human fib
4	76	100.0	760	2	Aaw27438 Human fib
5	76	100.0	760	6	Adi47452 Breast ca
6	76	100.0	760	7	Adn95552 Human BEC
7	76	100.0	760	8	Adi21351 Human sof
8	76	100.0	760	9	Adi14775 Tumor-agg
9	76	100.0	760	9	Aeb94159 Human wil
10	64	84.2	734	9	Aeb94218 Human sol
11	52	68.4	50	9	Aeb94178 Mouse sol
12	52	68.4	761	9	Aeb94163 Mouse wil
13	42	55.3	407	4	Aab97649 Human spe
14	42	55.3	425	8	Adi13246 Human mal
15	42	55.3	425	9	Aeal15133 Human pol
16	42	55.3	837	7	Abm86151 Rice abio
17	41	53.9	119	4	Aau45568 Propionib
18	41	53.9	119	6	Abm42087 Propionib
19	41	53.9	458	8	Ady08853 Plant ful
20	41	53.9	499	6	Abu20581 Protein e
21	41	53.9	642	5	Aam50586 Benign pr
22	41	53.9	666	4	Aam41498 Human pol
23	41	53.9	835	3	Aag31412 Arabidops
24	41	53.9	929	3	Aag31411 Arabidops

25	41	53.9	937	3	AAG31410	Arabidops
26	41	53.9	937	6	ABP59344	AT4G10590
27	41	53.9	2828	5	ABG31322	Human adl
28	41	53.9	2828	5	ABG31319	Human adl
29	41	53.9	2828	5	ABG32897	Human Adl
30	41	53.9	2828	5	AD117071	Human NOV
31	41	53.9	2828	5	AD117072	Human NOV
32	41	53.9	2828	6	ABJ37051	Human bre
33	41	53.9	2828	6	ABU56609	Lung can
34	41	53.9	2828	6	ABR47439	Breast ca
35	41	53.9	2828	7	ADJ81335	Human adl
36	41	53.9	2828	7	ADN39857	Cancer/an
37	41	53.9	2828	7	ADN38740	Cancer/an
38	41	53.9	2828	8	ADL02241	Human OCP
39	41	53.9	2828	8	ADQ17737	Human sof
40	41	53.9	2828	8	ADR67221	Human bia
41	41	53.9	2828	9	ADZ09884	Human bre
42	41	53.9	2828	9	AEA04479	Human pro
43	41	53.9	2853	6	ABR40110	Human cel
44	41	53.9	2993	4	ABG23884	Novel hum
45	41	53.9	5266	4	ABG08561	Novel hum
46	40	52.6	55	5	ABP11166	Human ORF
47	40	52.6	90	7	ADL33594	Transcrip
48	40	52.6	90	8	ADM72635	UME6 prot
49	40	52.6	90	8	ADP71011	Novel zin
50	40	52.6	90	9	ADV39071	Yeast UME
51	40	52.6	115	4	AAG91827	C glutami
52	40	52.6	210	7	ADL33681	Zinc fing
53	40	52.6	210	8	ADM72637	Amino aci
54	40	52.6	238	7	ADL33682	Zinc fing
55	40	52.6	238	7	ADL33684	Zinc fing
56	40	52.6	238	7	ADL33683	Zinc fing
57	40	52.6	238	8	ADM72638	Amino aci
58	40	52.6	238	8	ADM72639	Amino aci
59	40	52.6	238	8	ADM72640	Amino aci
60	40	52.6	310	8	ADR10155	Human pro
61	40	52.6	313	8	ADY24955	Plant ful
62	40	52.6	494	3	AAG40439	Arabidops
63	40	52.6	507	3	AAG40438	Arabidops
64	40	52.6	520	3	AAG40437	Arabidops
65	40	52.6	836	5	ABP35621	Fungal 2B
66	40	52.6	836	6	ABR53516	Protein s
67	40	52.6	836	7	ABK64454	Disease t
68	40	52.6	975	4	ABB71176	Drosophil
69	40	52.6	1584	2	AAY33727	Phototrab
70	39	51.3	13	9	ADM99256	Pig 12-LQ
71	39	51.3	142	8	ADN04796	Antipsori
72	39	51.3	142	8	ABM81456	Tumour as
73	39	51.3	142	8	ADP55198	Human PRO
74	39	51.3	142	9	ADY17301	PRO poly
75	39	51.3	144	2	AAY60106	Human end
76	39	51.3	144	8	ADR66428	Human pro
77	39	51.3	144	8	ADR66086	Human pro
78	39	51.3	144	8	ADR66353	Human pro
79	39	51.3	144	8	ADR66695	Human pro
80	39	51.3	277	3	AAG53675	Arabidops
81	39	51.3	333	3	AAG07543	Arabidops
82	39	51.3	333	3	AAG07542	Arabidops
83	39	51.3	361	3	AAG53674	Arabidops
84	39	51.3	361	3	AAG53673	Arabidops
85	39	51.3	362	3	AAG07541	Arabidops
86	39	51.3	421	4	ABG28196	Novel hum
87	39	51.3	497	4	ABG66268	Murine TA
88	39	51.3	630	8	ADO28948	Mouse nov
89	39	51.3	917	3	AAG38836	Arabidops
90	39	51.3	917	8	ADN72373	Thale cre
91	38	50.0	148	7	ADM04204	Human pro
92	38	50.0	183	4	ABM16106	Peptide #
93	38	50.0	183	4	ABB35091	Peptide #
94	38	50.0	183	4	ABB28597	Peptide #
95	38	50.0	183	4	ABB29913	Peptide #
96	38	50.0	183	4	ABB20510	Protein #
97	38	50.0	183	4	AAM68283	Human bon

88 38 50.0 183 4 RAM55913 Aam55913 Human bra
89 38 50.0 183 4 ABG49941 Abg49941 Human liv
100 38 50.0 183 4 AAM03833 Aam03833 Peptide #
101 38 50.0 183 5 ABG37824 Abg37824 Human pep
102 38 50.0 198 5 ABG77283 Abg77283 Selected
103 38 50.0 198 5 ABJ11185 Abj11185 Yeast sel
104 38 50.0 232 4 AAU36976 Aau36976 Staphyloc
105 38 50.0 232 4 AAU34064 Aau34064 Staphyloc
106 38 50.0 259 6 ABR53507 Abr53507 Protein s
107 38 50.0 259 7 ADK62586 Adk62586 Disease t
108 38 50.0 308 4 ABG03387 Abg03387 Novel hum
109 38 50.0 312 8 ADY05395 Ady05395 Plant ful
110 38 50.0 347 8 ADY08738 Ady08738 Plant ful
111 38 50.0 368 8 ADS43555 Ads43555 Bacterial
112 38 50.0 371 2 RAY08963 Ray08963 A. gossyp
113 38 50.0 495 4 ABB59762 Abb59762 Drosophil
114 38 50.0 549 7 ABM87265 Abm87265 Rice abio
115 38 50.0 798 4 ABG19506 Abg19506 Novel hum
116 38 50.0 966 8 ADN19519 Adn19519 Bacterial
117 38 50.0 1278 3 AAB18277 Aab18277 Plasmodi
118 38 50.0 2652 5 ABG31321 Abg31321 Predicted
119 38 50.0 2652 8 ADL02250 Adl02250 Human OCP
120 38 50.0 2845 5 ABP70049 Abp70049 Human NOV
121 37.5 49.3 710 7 ABO71987 Abo71987 Pseudomon
122 37 48.7 40 7 ABB81089 Abb81089 Hrp secre
123 37 48.7 74 4 ABB68279 Abb68279 Drosophil
124 37 48.7 83 3 AAY58544 Aay58544 Human som
125 37 48.7 123 4 AAU27526 Aau27526 Human G-P
126 37 48.7 198 7 ADB76857 Adb76857 A. gossyp
127 37 48.7 219 8 ADT59464 Adt59464 Plant pol
128 37 48.7 228 3 AAG42418 Aag42418 Arabidops
129 37 48.7 229 3 AAG42416 Aag42416 Arabidops
130 37 48.7 241 3 AAG42416 Aag42416 Arabidops
131 37 48.7 301 8 ADN23805 Adn23805 Bacterial
132 37 48.7 304 5 ADI17191 Adi17191 Human NOV
133 37 48.7 340 4 AAB69185 Aab69185 Human his
134 37 48.7 345 2 AAW09405 Aaw09405 Pineal gl
135 37 48.7 345 5 ABB90207 Abb90207 Human pol
136 37 48.7 345 6 ABG73318 Abg73318 Human pin
137 37 48.7 345 7 ADF69096 Adf69096 Human MP5
138 37 48.7 358 8 ADW44172 Adw44172 msra-1 pr
139 37 48.7 366 3 AAY93266 Aay93266 Amino aci
140 37 48.7 375 7 ABO61153 Abo61153 Klebsiell
141 37 48.7 405 6 ABU20572 Abu20572 Protein e
142 37 48.7 408 7 ADC21265 Adc21265 Plasmodi
143 37 48.7 414 7 ABO66061 Abo66061 Klebsiell
144 37 48.7 418 2 AAR39263 Aar39263 Human som
145 37 48.7 418 4 AAB99053 Aab99053 Human som
146 37 48.7 418 4 ABB56390 Abb56390 Non-endog
147 37 48.7 418 6 ABB81915 Abb81915 Human som
148 37 48.7 418 7 AAE39512 Aae39512 Human GPC
149 37 48.7 418 8 ADW42067 Adw42067 Human Som
150 37 48.7 418 8 ADO29650 Ado29650 Human GPC
151 37 48.7 418 8 ADL15201 Adl15201 Human som
152 37 48.7 418 8 ADR40551 Adr40551 Human som
153 37 48.7 458 8 ADN21257 Adn21257 Bacterial
154 37 48.7 477 5 ABP65751 Abp65751 Bifidobac
155 37 48.7 481 8 ADP22602 Adp22602 Sea-squir
156 37 48.7 544 8 ADN19723 Adn19723 Bacterial
157 37 48.7 575 7 ABM89158 Abm89158 Rice abio
158 37 48.7 737 6 ABU50610 Abu50610 Protein e
159 37 48.7 805 8 ADT56254 Adt56254 Plant pol
160 37 48.7 996 8 ADT49895 Adt49895 Murine BC
161 37 48.7 1021 7 ABO78659 Abo78659 Pseudomon
162 37 48.7 1354 4 ABB62359 Abb62359 Drosophil
163 37 48.7 4498 4 ABB58595 Abb58595 Drosophil
164 36.5 48.0 355 5 ABB51368 Abb51368 Human MDD
165 36.5 48.0 601 6 AAO16428 Aao16428 Human nuc
166 36.5 48.0 601 7 ADI60222 Adi60222 Secreted
167 36.5 48.0 637 8 ADG67577 Adg67577 Novel hum
168 36.5 48.0 651 6 ABU11549 Abu11549 Human MDD
169 36.5 48.0 673 6 ABU11692 Abu11692 Human MDD
170 36.5 48.0 678 8 ADN72297 Adn72297 Thale cre

171 36 47.4 26 2 AAR47423 Aar47423 PDGF-acti
172 36 47.4 141 2 AAU30879 Aau30879 Novel humo
173 36 47.4 163 2 AAW80707 Aaw80707 S. pneumo
174 36 47.4 173 8 ADS22771 Ads22771 Bacterial
175 36 47.4 191 2 AAW38584 Aaw38584 Streptoco
176 36 47.4 194 3 AAB28208 Aab28208 Novel hum
177 36 47.4 205 8 ADI27148 Adi27148 Mouse LRP
178 36 47.4 211 8 ABO60418 Abo60418 Human gen
179 36 47.4 219 7 ABO80770 Abo80770 Pseudomon
180 36 47.4 221 4 ABG29021 Abg29021 Novel hum
181 36 47.4 242 4 AAU37458 Aau37458 Staphyloc
182 36 47.4 243 7 ABM89054 Abm89054 Rice abio
183 36 47.4 251 4 AAU37124 Aau37124 Staphyloc
184 36 47.4 252 5 ABG77239 Abg77239 Selected
185 36 47.4 252 5 ABJ11110 Abj11110 Yeast sel
186 36 47.4 269 7 ADH87367 Adh87367 Enterococ
187 36 47.4 277 6 AAU50423 Aau50423 Propionib
188 36 47.4 277 6 ABM46942 Abm46942 Propionib
189 36 47.4 279 8 ADI10339 Adi10339 Human pro
190 36 47.4 282 3 AAY94436 Aay94436 Streptoco
191 36 47.4 283 6 ABM73470 Abm73470 Staphyloc
192 36 47.4 295 4 AAU38098 Aau38098 Streptoco
193 36 47.4 295 4 AAU37758 Aau37758 Streptoco
194 36 47.4 295 6 ABU01332 Abu01332 S. pneumo
195 36 47.4 295 8 ADK48621 Adk48621 Streptoco
196 36 47.4 295 8 ADR94735 Adr94735 Novel S.
197 36 47.4 295 9 AEA58605 Aea58605 Streptoco
198 36 47.4 311 8 ADS45246 Ads45246 Bacterial
199 36 47.4 407 8 ADN20853 Adn20853 Bacterial
200 36 47.4 414 6 ABM65766 Abm65766 Propionib
201 36 47.4 420 2 AAW50282 Aaw50282 Canine he
202 36 47.4 423 4 AAU34377 Aau34377 Staphyloc
203 36 47.4 424 3 AAG27964 Aag27964 Arabidops
204 36 47.4 441 5 AAM47938 Aam47938 Human zin
205 36 47.4 441 6 AAE33785 Aae33785 Human nuc
206 36 47.4 441 6 ADA55416 Ada55416 Human pro
207 36 47.4 501 5 ABB09170 Abb09170 Methylomo
208 36 47.4 501 5 ABG61555 Abg61555 High grow
209 36 47.4 551 6 ABU25027 Abu25027 Protein e
210 36 47.4 552 6 ABU23845 Abu23845 Protein e
211 36 47.4 579 6 ABU16559 Abu16559 Protein e
212 36 47.4 592 4 ABG24277 Abg24277 Novel hum
213 36 47.4 606 6 ADA55496 Ada55496 Human pro
214 36 47.4 613 4 ABG25954 Abg25954 Novel hum
215 36 47.4 620 6 ABR53180 Abr53180 Protein s
216 36 47.4 620 7 ADK63212 Adk63212 Disease t
217 36 47.4 627 4 ABB61475 Abb61475 Drosophil
218 36 47.4 645 8 ADL04397 Adl04397 M. catarr
219 36 47.4 733 4 AAG70721 Aag70721 S. cerevis
220 36 47.4 1066 8 ADS44238 Ads44238 Bacterial
221 36 47.4 1089 7 ABO71588 Abo71588 Pseudomon
222 36 47.4 1250 8 ADN19019 Adn19019 Bacterial
223 36 47.4 1679 6 ABR52863 Abr52863 Protein s
224 36 47.4 1679 7 ADK62332 Adk62332 Disease t
225 36 47.4 2162 4 ABB64574 Abb64574 Drosophil
226 36 47.4 2162 8 ABO07970 Abo07970 Fly polyp
227 35.5 46.7 179 4 ABB14761 Abb14761 Human ner
228 35.5 46.7 303 7 ADM25479 Adm25479 Hyperther
229 35.5 46.7 683 4 AAY97729 Aay97729 NS protei
230 35.5 46.7 683 5 AAE22896 Aae22896 Rhesus ma
231 35.5 46.7 683 5 AAE28653 Aae28653 Rhesus ma
232 35.5 46.7 683 6 ABU64874 Abu64874 NS protei
233 35.5 46.7 683 6 ABU64769 Abu64769 Farroviru
234 35.5 46.7 683 6 ABR43407 Abr43407 Rhesus ma
235 35.5 46.7 683 7 ADI40298 Adi40298 Rhesus ma
236 35.5 46.7 1467 7 ABO68911 Abo68911 Pseudomon
237 35.5 46.7 13 9 ADM99247 Adm99247 Mouse 12-
238 35 46.1 14 4 AAW96902 Aaw96902 Human pep
239 35 46.1 55 5 ABP31037 Abp31037 Human ORF
240 35 46.1 62 4 AAU57957 Aau57957 Propionib
241 35 46.1 62 6 ABM54476 Abm54476 Propionib
242 35 46.1 67 4 ABG09590 Abg09590 Novel hum

244	35	46.1	72	3	AAG02795	Aag02795 Human sec	317	35	46.1	373	8	ADY24582	Ady24582 Plant ful
245	35	46.1	73	3	AAy65124	AAy65124 Human 5'	318	35	46.1	374	2	AAW37497	Aaw37497 Human TMP
246	35	46.1	73	4	ABG17830	ABG17830 Novel hum	319	35	46.1	374	2	AAW07663	Aaw07663 Human tra
247	35	46.1	73	8	ADU72688	ADU72688 Signal pe	320	35	46.1	374	3	AAy51814	AAy51814 Human emb
248	35	46.1	73	9	ADZ73679	ADZ73679 Human inc	321	35	46.1	374	3	AAy51814	AAy51814 Human emb
249	35	46.1	83	3	AAy58543	AAy58543 Mouse som	322	35	46.1	374	4	AAm93196	AAm93196 Human pol
250	35	46.1	83	3	AAy58542	AAy58542 Rat somat	323	35	46.1	374	4	AAm93196	AAm93196 Human pol
251	35	46.1	84	4	AAU47430	AAU47430 Propionib	324	35	46.1	374	4	AAU29226	AAU29226 Human PRO
252	35	46.1	84	6	ABM43949	ABM43949 Propionib	325	35	46.1	374	5	ABB90330	ABB90330 Human pol
253	35	46.1	88	5	ABM43949	ABM43949 Human sec	326	35	46.1	374	5	ABG61836	ABG61836 Prostate
254	35	46.1	88	5	ABM43949	ABM43949 Human sec	327	35	46.1	374	5	ABG61836	ABG61836 Prostate
255	35	46.1	88	5	ABM43949	ABM43949 Human alb	328	35	46.1	374	5	ABU76537	ABU76537 Tumour-as
256	35	46.1	88	5	ABM43949	ABM43949 Human alb	329	35	46.1	374	5	ABU76537	ABU76537 Tumour-as
257	35	46.1	88	8	ADL78418	ADL78418 Albumin f	330	35	46.1	374	6	ABU88150	ABU88150 Novel hum
258	35	46.1	88	8	ADL78418	ADL78418 Albumin f	331	35	46.1	374	6	ABU88150	ABU88150 Novel hum
259	35	46.1	91	7	ABO75194	ABO75194 Pseudomon	332	35	46.1	374	6	ABU88150	ABU88150 Novel hum
260	35	46.1	103	4	AAU58156	AAU58156 Propionib	333	35	46.1	374	6	ABU88150	ABU88150 Novel hum
261	35	46.1	103	6	ABM54675	ABM54675 Propionib	334	35	46.1	374	6	ABU88150	ABU88150 Novel hum
262	35	46.1	105	9	ABE39638	ABE39638 L. pneumo	335	35	46.1	374	6	ABU88150	ABU88150 Novel hum
263	35	46.1	108	4	ABG01669	ABG01669 Novel hum	336	35	46.1	374	6	ABU88150	ABU88150 Novel hum
264	35	46.1	111	9	ABE36216	ABE36216 L. pneumo	337	35	46.1	374	6	ABU88150	ABU88150 Novel hum
265	35	46.1	112	4	AAU64288	AAU64288 Propionib	338	35	46.1	374	6	ABU88150	ABU88150 Novel hum
266	35	46.1	112	6	ABM60807	ABM60807 Propionib	339	35	46.1	374	6	ABU88150	ABU88150 Novel hum
267	35	46.1	117	5	ABM60807	ABM60807 Human sec	340	35	46.1	374	6	ABU88150	ABU88150 Novel hum
268	35	46.1	117	5	ABM60807	ABM60807 Human sec	341	35	46.1	374	6	ABU88150	ABU88150 Novel hum
269	35	46.1	117	5	ABM60807	ABM60807 Human sec	342	35	46.1	374	6	ABU88150	ABU88150 Novel hum
270	35	46.1	121	8	ADL78416	ADL78416 Albumin f	343	35	46.1	374	6	ABU88150	ABU88150 Novel hum
271	35	46.1	124	8	ADL78416	ADL78416 Albumin f	344	35	46.1	374	6	ABU88150	ABU88150 Novel hum
272	35	46.1	133	3	ABG33822	ABG33822 Novel hum	345	35	46.1	374	6	ABU88150	ABU88150 Novel hum
273	35	46.1	138	4	ABG05104	ABG05104 Novel hum	346	35	46.1	374	6	ABU88150	ABU88150 Novel hum
274	35	46.1	138	4	ABG14739	ABG14739 Novel hum	347	35	46.1	374	6	ABU88150	ABU88150 Novel hum
275	35	46.1	138	4	ABG14739	ABG14739 Novel hum	348	35	46.1	374	6	ABU88150	ABU88150 Novel hum
276	35	46.1	160	9	ABE41493	ABE41493 L. pneumo	349	35	46.1	374	6	ABU88150	ABU88150 Novel hum
277	35	46.1	161	9	ABE38206	ABE38206 L. pneumo	350	35	46.1	374	6	ABU88150	ABU88150 Novel hum
278	35	46.1	206	1	ABG08223	ABG08223 Novel hum	351	35	46.1	374	6	ABU88150	ABU88150 Novel hum
279	35	46.1	210	7	ABM02585	ABM02585 Streptoco	352	35	46.1	374	6	ABU88150	ABU88150 Novel hum
280	35	46.1	211	2	AAW14547	AAW14547 Streptoco	353	35	46.1	374	6	ABU88150	ABU88150 Novel hum
281	35	46.1	213	4	ABU61139	ABU61139 Protein e	354	35	46.1	374	6	ABU88150	ABU88150 Novel hum
282	35	46.1	232	6	ABM70990	ABM70990 Staphyloc	355	35	46.1	374	6	ABU88150	ABU88150 Novel hum
283	35	46.1	232	6	ABM70990	ABM70990 Staphyloc	356	35	46.1	374	6	ABU88150	ABU88150 Novel hum
284	35	46.1	237	4	AAW14149	AAW14149 Peptide #	357	35	46.1	374	6	ABU88150	ABU88150 Novel hum
285	35	46.1	237	4	ABE33094	ABE33094 Peptide #	358	35	46.1	374	6	ABU88150	ABU88150 Novel hum
286	35	46.1	237	4	AAW26655	AAW26655 Peptide #	359	35	46.1	374	6	ABU88150	ABU88150 Novel hum
287	35	46.1	237	4	ABE27922	ABE27922 Human pep	360	35	46.1	374	6	ABU88150	ABU88150 Novel hum
288	35	46.1	237	4	ABE18562	ABE18562 Protein #	361	35	46.1	374	6	ABU88150	ABU88150 Novel hum
289	35	46.1	237	4	AAW66279	AAW66279 Human bon	362	35	46.1	374	6	ABU88150	ABU88150 Novel hum
290	35	46.1	237	4	AAW53891	AAW53891 Human bon	363	35	46.1	374	6	ABU88150	ABU88150 Novel hum
291	35	46.1	237	4	ABG47943	ABG47943 Human liv	364	35	46.1	374	6	ABU88150	ABU88150 Novel hum
292	35	46.1	237	4	AAW01887	AAW01887 Peptide #	365	35	46.1	374	6	ABU88150	ABU88150 Novel hum
293	35	46.1	237	5	ABG35926	ABG35926 Human pep	366	35	46.1	374	6	ABU88150	ABU88150 Novel hum
294	35	46.1	270	6	ADW55394	ADW55394 Human pro	367	35	46.1	374	6	ABU88150	ABU88150 Novel hum
295	35	46.1	274	4	AAG72403	AAG72403 Human OR-	368	35	46.1	374	6	ABU88150	ABU88150 Novel hum
296	35	46.1	274	4	AAG71655	AAG71655 Human olf	369	35	46.1	374	6	ABU88150	ABU88150 Novel hum
297	35	46.1	283	5	ABP27535	ABP27535 Streptoco	370	35	46.1	374	6	ABU88150	ABU88150 Novel hum
298	35	46.1	283	5	ADW81411	ADW81411 Streptoco	371	35	46.1	374	6	ABU88150	ABU88150 Novel hum
299	35	46.1	301	9	ABM95819	ABM95819 M. xanthu	372	35	46.1	374	6	ABU88150	ABU88150 Novel hum
300	35	46.1	304	8	ADW87968	ADW87968 Streptoco	373	35	46.1	374	6	ABU88150	ABU88150 Novel hum
301	35	46.1	304	8	ADW79221	ADW79221 Streptoco	374	35	46.1	374	6	ABU88150	ABU88150 Novel hum
302	35	46.1	314	8	ADW43752	ADW43752 Bacterial	375	35	46.1	374	6	ABU88150	ABU88150 Novel hum
303	35	46.1	321	8	ADY05465	ADY05465 Plant ful	376	35	46.1	374	6	ABU88150	ABU88150 Novel hum
304	35	46.1	336	4	AAU69963	AAU69963 Human pro	377	35	46.1	374	6	ABU88150	ABU88150 Novel hum
305	35	46.1	336	4	ABU71854	ABU71854 Prostate	378	35	46.1	374	6	ABU88150	ABU88150 Novel hum
306	35	46.1	336	6	ABR54535	ABR54535 Prostate	379	35	46.1	374	6	ABU88150	ABU88150 Novel hum
307	35	46.1	336	7	ADW14390	ADW14390 Human pro	380	35	46.1	374	6	ABU88150	ABU88150 Novel hum
308	35	46.1	336	7	ADW26806	ADW26806 Human pro	381	35	46.1	374	6	ABU88150	ABU88150 Novel hum
309	35	46.1	347	4	AAW01225	AAW01225 Human pro	382	35	46.1	374	6	ABU88150	ABU88150 Novel hum
310	35	46.1	347	4	AAU69870	AAU69870 Human pro	383	35	46.1	374	6	ABU88150	ABU88150 Novel hum
311	35	46.1	347	4	ABU71761	ABU71761 Prostate	384	35	46.1	374	6	ABU88150	ABU88150 Novel hum
312	35	46.1	347	5	ABR95330	ABR95330 Human P76	385	35	46.1	374	6	ABU88150	ABU88150 Novel hum
313	35	46.1	347	5	ABR95330	ABR95330 Human P76	386	35	46.1	374	6	ABU88150	ABU88150 Novel hum
314	35	46.1	347	6	ADW14040	ADW14040 Prostate	387	35	46.1	374	6	ABU88150	ABU88150 Novel hum
315	35	46.1	347	7	ADG26456	ADG26456 Human pro	388	35	46.1	374	6	ABU88150	ABU88150 Novel hum
316	35	46.1	364	8	ABO84719	ABO84719 Human can	389	35	46.1	374	6	ABU88150	ABU88150 Novel hum

390	35	46.1	374	6	ABR65361	Human sec
391	35	46.1	374	6	ABR65883	Human sec
392	35	46.1	374	6	ABR71995	Human sec
393	35	46.1	374	6	ABU85475	Human PRO
394	35	46.1	374	6	ABU89165	Human sec
395	35	46.1	374	6	ABU83245	Human sec
396	35	46.1	374	6	ABU95101	Novel hum
397	35	46.1	374	6	ABU90649	Novel hum
398	35	46.1	374	6	ABU84160	Human sec
399	35	46.1	374	6	ABU93811	Novel hum
400	35	46.1	374	6	ABR65056	Human sec
401	35	46.1	374	6	ABR68888	Human sec
402	35	46.1	374	6	ABO06704	Human sec
403	35	46.1	374	6	ABR93249	Human sec
404	35	46.1	374	6	ABU57133	Human PRO
405	35	46.1	374	6	ABU86085	Novel hum
406	35	46.1	374	6	ABU82372	Novel hum
407	35	46.1	374	6	ABU87383	Human PRO
408	35	46.1	374	6	ABU83855	Human sec
409	35	46.1	374	6	ABO08229	Human PRO
410	35	46.1	374	6	ABU81940	Novel hum
411	35	46.1	374	6	ABU66104	Novel hum
412	35	46.1	374	6	ABR59933	Human sec
413	35	46.1	374	6	ABU94121	Novel hum
414	35	46.1	374	6	ABU99974	Human sec
415	35	46.1	374	6	ABR66644	Human sec
416	35	46.1	374	6	ABR91062	Human sec
417	35	46.1	374	6	ABU94489	Human PRO
418	35	46.1	374	6	ABU79371	Human PRO
419	35	46.1	374	6	ABU86700	Human sec
420	35	46.1	374	6	ABU87005	Novel hum
421	35	46.1	374	6	ABU94794	Human PRO
422	35	46.1	374	6	ABO04721	Human PRO
423	35	46.1	374	6	ABR70470	Human sec
424	35	46.1	374	6	ABU98635	Human PRO
425	35	46.1	374	6	ABR66034	Human sec
426	35	46.1	374	6	ABR64751	Human sec
427	35	46.1	374	6	ABU79676	Human PRO
428	35	46.1	374	6	ABU93067	Human sec
429	35	46.1	374	6	ABU96026	Human PRO
430	35	46.1	374	6	ABU91246	Novel hum
431	35	46.1	374	6	ABU90339	Novel hum
432	35	46.1	374	6	ABO09754	Human sec
433	35	46.1	374	6	ABO11026	Human sec
434	35	46.1	374	6	ABR71080	Human sec
435	35	46.1	374	6	ABU87688	Human PRO
436	35	46.1	374	6	ABU91556	Human PRO
437	35	46.1	374	6	ABU84770	Human sec
438	35	46.1	374	6	ABR69860	Human sec
439	35	46.1	374	6	ABU80237	Human PRO
440	35	46.1	374	6	ABU93506	Human PRO
441	35	46.1	374	6	ABO10059	Human sec
442	35	46.1	374	6	ABO09144	Human sec
443	35	46.1	374	6	ABU10712	Human sec
444	35	46.1	374	6	ABU95721	Human PRO
445	35	46.1	374	6	ABU96930	Novel hum
446	35	46.1	374	6	ABR70775	Human sec
447	35	46.1	374	6	ABO05126	Novel hum
448	35	46.1	374	6	ABO08534	Human sec
449	35	46.1	374	6	ABO05741	Human sec
450	35	46.1	374	6	ABR74130	Human sec
451	35	46.1	374	6	ABR95722	Human sec
452	35	46.1	374	6	ABR81019	Human sec
453	35	46.1	374	6	ABR81324	Human sec
454	35	46.1	374	6	ABM01020	Human sec
455	35	46.1	374	6	ABR88622	Human sec
456	35	46.1	374	6	ABM77443	Human sec
457	35	46.1	374	6	ABO28927	Human sec
458	35	46.1	374	6	ABO31672	Human sec
459	35	46.1	374	6	ABM08089	Human sec
460	35	46.1	374	6	ABO40569	Human sec
461	35	46.1	374	6	ABO35994	Human PRO
462	35	46.1	374	6	ABO44133	Human PRO

463	35	46.1	374	6	ADA78158	Human sec
464	35	46.1	374	6	ABM24928	Human sec
465	35	46.1	374	6	ABO03196	Human sec
466	35	46.1	374	6	ABR90452	Human sec
467	35	46.1	374	6	ABM17366	Human sec
468	35	46.1	374	6	ABR95112	Human sec
469	35	46.1	374	6	ABR95417	Human sec
470	35	46.1	374	6	ABO21655	Human sec
471	35	46.1	374	6	ABR97919	Human sec
472	35	46.1	374	6	ABR87707	Human sec
473	35	46.1	374	6	ABM77748	Human sec
474	35	46.1	374	6	ABM27978	Human sec
475	35	46.1	374	6	ABM06259	Human sec
476	35	46.1	374	6	ABM03765	Human sec
477	35	46.1	374	6	ABM26453	Human sec
478	35	46.1	374	6	ABO48235	Human sec
479	35	46.1	374	6	ABR92977	Human sec
480	35	46.1	374	6	ABO24738	Human sec
481	35	46.1	374	6	ABM11749	Human sec
482	35	46.1	374	6	ABM02850	Human sec
483	35	46.1	374	6	ABM16146	Human sec
484	35	46.1	374	6	ABO27707	Human sec
485	35	46.1	374	6	ABM29198	Human sec
486	35	46.1	374	6	ABM07174	Human sec
487	35	46.1	374	6	ABM21268	Human sec
488	35	46.1	374	6	ABM09614	Human sec
489	35	46.1	374	6	ABO41484	Human sec
490	35	46.1	374	6	ABO36299	Human PRO
491	35	46.1	374	6	ABO43828	Human sec
492	35	46.1	374	6	ABM76528	Human sec
493	35	46.1	374	6	ABM76224	Human sec
494	35	46.1	374	6	ABM25843	Human sec
495	35	46.1	374	6	ABM25843	Human sec
496	35	46.1	374	6	ABM26148	Human sec
497	35	46.1	374	6	ABO03501	Human sec
498	35	46.1	374	6	ABO02586	Human sec
499	35	46.1	374	6	ABR90757	Human sec
500	35	46.1	374	6	ABR73825	Human sec

ALIGNMENTS

RESULT 1

ADR47257
ID ADR47257 standard; peptide; 15 AA.

XX AC ADR47257;

XX DT 18-NOV-2004 (first entry)

XX DE Human alpha2-antiplasmin cleaving enzyme N-terminal sequence.

XX antiinflammatory; cytostatic; vulnery; antiarteriosclerotic;
XX antithrombotic; vascular; cerebroprotective; pulmonary;
XX fibrin-related disorder; inflammatory condition; arthritis;
XX organ fibrosis; scarring; cancer; metastasis; atherothrombotic disease;
XX coronary artery thrombosis; stroke; pulmonary embolism; thrombosis;
XX human; N-terminus; alpha2-antiplasmin cleaving enzyme.

OS Homo sapiens.

XX WO2004072240-A2.

XX 26-AUG-2004.

XX 07-FEB-2004; 2004WO-US003398.

XX 07-FEB-2003; 2003US-0445774P.

XX 06-FEB-2004; 2004US-00774242.

XX (MCKE/) MCKEE P A.

XX (LEEK/) LEE K N.

PA (JACK/) JACKSON K W.
 PA (CHRI/) CHRISTIANSEN V J.
 PI McKeen PA, Lee KN, Jackson KW, Christiansen VJ;
 XX WPI; 2004-625848/60.
 DR
 XX
 XX New alpha-2-antiplasmin cleaving enzyme, useful for treating conditions
 PT involving fibrin, e.g. inflammatory conditions such as all forms of
 PT arthritis, organ fibrosis, undesirable scarring, cancer, or
 PT atherothrombotic disease.
 XX
 PS Claim 2; SEQ ID NO 1; 40pp; English.
 XX
 CC The invention relates to an alpha2-antiplasmin cleaving enzyme comprising
 CC a protein having a molecular weight of 180 kD in a dimeric form as
 CC determined by SDS-PAGE, where each subunit of the dimeric form has a
 CC molecular weight of 97 kD as determined by SDS-PAGE, and where the enzyme
 CC cleaves precursor alpha2-antiplasmin at the Pro12-Asn13 bond. The enzyme,
 CC inhibitors of the enzyme and methods for detecting such inhibitors are
 CC useful for treating conditions involving fibrin, e.g. inflammatory
 CC conditions such as all forms of arthritis, organ fibrosis, undesirable
 CC scarring, cancer or its metastases; or atherothrombotic disease such as
 CC coronary artery thrombosis, stroke, pulmonary embolism, all other forms
 CC of arterial and venous thromboses. This peptide corresponds to the N-
 CC terminal sequence of the alpha2-antiplasmin cleaving enzyme.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 76; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IVLRPSRVHNSSENT 15
 DB 1 IVLRPSRVHNSSENT 15
 RESULT 2
 AEB94176
 ID AEB94176 standard; protein; 50 AA.
 XX
 AC AEB94176;
 XX
 DT 06-OCT-2005 (first entry)
 XX
 DE Human soluble FAP alpha dimer N-terminal fragment SEQ ID NO:19.
 XX
 KW immune inhibition; fibroblast activation protein alpha dimer;
 KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
 KW immune disorder; neurological disease; autoimmune disease;
 KW immunosuppressive; graft versus host disease; transplant rejection;
 KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
 KW musculoskeletal disease; allergy; antiasthmatic; asthma; antiasthmatic;
 KW inflammation; respiratory disease; atherosclerosis; arteriosclerotic;
 KW cardiovascular disease; metabolic disorder; hashimoto disease;
 KW antithyroid; endocrine disease; inflammatory bowel disease;
 KW antinflammatory; gastrointestinal-gen.; gastrointestinal disease;
 KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
 KW autoimmune hepatitis; antinflammatory; hepatotropic;
 KW systemic lupus erythematosus; dermatological; dermatologic disease;
 KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
 KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
 KW psoriasis; antipsoriatic; graves disease; antithyroid;
 KW respiratory syncytial virus infection; respiratory-gen.; virucide.
 XX
 OS Homo sapiens.
 XX
 PN WO2005071073-A1.
 XX
 PD 04-AUG-2005.
 XX
 PF 10-JAN-2005; 2005NO-US000709.
 XX
 XX 09-JAN-2004; 2004US-0535577P.
 XX (POIN-) POINT THERAPEUTICS INC.
 XX McLean PA, Jones B, Miller GT, Jesson MI;
 XX WPI; 2005-564220/57.
 DR
 XX
 XX Down-regulating an immune response comprises administering to a subject
 PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
 PT amount effective to down-regulate an immune response.
 XX
 PS Example 1.4; SEQ ID NO 19; 177pp; English.
 XX
 CC The invention relates to a method of down-regulating an immune response,
 CC which comprises administering to a subject a fibroblast activation
 CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
 CC an immune response. Also included are the following: a composition
 CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
 CC carrier, where the composition is sterile and lacks an adjuvant; a
 CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
 CC acceptable carrier, and a non-adjuvant second agent; a composition
 CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
 CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
 CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
 CC The method further comprises administering to the subject a second agent.
 CC The second agent is an anti-inflammatory agent, immunosuppressant, or
 CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
 CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
 CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
 CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
 CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
 CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
 CC substitution relative to wild type FAP alpha dimer. The amino acid
 CC substitution is present in the beta-propeller domain, the catalytic
 CC domain, or an N-linked glycosylation site and alters disulfide bond
 CC formation. The immune response is an especially an IL-1 mediated
 CC condition, abnormal immune response selected from inflammation,
 CC autoimmune disease, sepsis, graft versus host disease, transplant
 CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
 CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
 CC response is subsequent to an infection, such as an RSV infection. The
 CC autoimmune disease is selected from C, autoimmune thyroiditis, systemic
 CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
 CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
 CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
 CC sclerosis. The subject does not have cancer or a predisposition to
 CC cancer. The present sequence represents the N-terminal 50 amino acids of
 CC human wild type FAP.
 XX
 SQ Sequence 50 AA;
 Query Match 100.0%; Score 76; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 4.6e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IVLRPSRVHNSSENT 15
 DB 24 IVLRPSRVHNSSENT 38
 RESULT 3
 AAW31963
 ID AAW31963 standard; protein; 759 AA.
 XX
 AC AAW31963;
 XX
 DT 21-JAN-1998 (first entry)
 XX
 DE Human fibroblast activation protein-alpha (FAP-alpha).
 XX
 KW Fibroblast activation protein-alpha; FAP-alpha; human;

type II integral membrane proteins; epithelial cancer.

OS Homo sapiens.

XX Key Location/Qualifiers

XX Protein 1..759

XX FT /label= FAP-alpha

XX FT /note= "Fibroblast activation protein alpha"

XX FT Modified-site 49..51

XX FT /note= "Asn is N-glycosylated"

XX FT Region 191..209

XX FT /label= fap-1

XX FT /note= "Putative functional element; not defined further

XX FT in specification"

XX FT Modified-site 214..216

XX FT /note= "Asn is N-glycosylated"

XX FT Region 220..240

XX FT /label= fap-2

XX FT /note= "Putative functional element; not defined further

XX FT in specification"

XX FT Region 510..521

XX FT /label= fap-3

XX FT /note= "Putative functional element; not defined further

XX FT in specification"

XX FT Region 625..668

XX FT /note= "alternative sequence found in AAW27438"

XX FT Modified-site 637..639

XX FT /note= "Asn is N-glycosylated"

XX FT Modified-site 678..680

XX FT /note= "Asn is N-glycosylated"

XX FT WO9734927-A1.

XX FT 25-SEP-1997.

XX FT 12-MAR-1997; 97WO-US004215.

XX FT 18-MAR-1996; 96US-00619280.

XX FT (LUDW-) LUDWIG INST CANCER RES.

XX FT (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX FT Zimmerman R, Park JE, Rettig W, Old LJ;

XX FT WPI; 1997-480164/44.

XX FT N-PSDB; AAT89286.

XX FT Dimeric fibroblast activation protein alpha - capable of degrading

XX FT extracellular matrix protein, useful to treat epithelial cancers.

XX FT Example 5; Fig 1; 40pp; English.

XX FT This sequence represents the dimeric fibroblast activation protein alpha

XX FT (FAP-alpha). The FAP-alpha molecule is capable of degrading extracellular

XX FT matrix proteins and exhibits structural features typical of type II

XX FT integral membrane proteins, including a large COOH-terminal extracellular

XX FT domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.

XX FT The putative extracellular domain is described as containing five

XX FT three segments corresponding to highly conserved catalytic domains

XX FT characteristic of serine proteases. Most common types of epithelial

XX FT cancers including breast, lung, skin, pancreas and colorectal carcinomas

XX FT contain abundant FAP-alpha reactive stromal fibroblasts which accompany

XX FT tumour blood vessels. The induction of FAP-alpha fibroblasts at times and

XX FT sites of tissue remodelling during foetal development, tissue repair and

XX FT carcinogenesis is consistent with its putative role in normal fibroblast

XX FT physiology. FAP alpha inhibitors, particularly a collagen derivative or

XX FT (S)-valyl-pyrrolidine-2(R)-boronic acid, may be used to treat various

XX FT epithelial cancers, e.g. breast, lung, skin, pancreas or colon cancer.

XX FT Note: This sequence differs from the amino acid sequence of FAP-alpha

XX FT found in AAW27438 between residues 625 to 670

XX FT Sequence 759 AA;

Query Match 100.0%; Score 76; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSENT 15

DB 24 IVLRPSRVHNSSENT 38

RESULT 4

AAW27438

ID AAW27438 standard; protein; 760 AA.

XX AC AAW27438;

XX XX 21-JAN-1998 (first entry)

XX DE Human fibroblast activation protein-alpha (FAP-alpha).

XX KW Fibroblast activation protein-alpha; FAP-alpha; human;

XX KW type II integral membrane proteins; epithelial cancer.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT 1..760

XX FT /label= FAP-alpha

XX FT /note= "Fibroblast activation protein alpha"

XX FT Region 625..670

XX FT /note= "alternative sequence found in AAW31963"

XX FT WO9734927-A1.

XX FT 25-SEP-1997.

XX FT 12-MAR-1997; 97WO-US004215.

XX FT 18-MAR-1996; 96US-00619280.

XX FT (LUDW-) LUDWIG INST CANCER RES.

XX FT (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX FT Zimmerman R, Park JE, Rettig W, Old LJ;

XX FT WPI; 1997-480164/44.

XX FT N-PSDB; AAT90148.

XX FT Dimeric fibroblast activation protein alpha - capable of degrading

XX FT extracellular matrix protein, useful to treat epithelial cancers.

XX FT Claim 2; Page 25-28; 40pp; English.

XX FT This sequence represents a dimeric fibroblast activation protein alpha

XX FT (FAP-alpha) which is capable of degrading extracellular matrix proteins.

XX FT The FAP-alpha molecule exhibits structural features typical of type II

XX FT integral membrane proteins, including a large COOH-terminal extracellular

XX FT domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.

XX FT Most common types of epithelial cancers including breast, lung, skin,

XX FT pancreas and colorectal carcinomas contain abundant FAP-alpha reactive

XX FT stromal fibroblasts which accompany tumour blood vessels. The induction

XX FT of FAP-alpha fibroblasts at times and sites of tissue remodelling during

XX FT fetal development, tissue repair and carcinogenesis is consistent with

XX FT its putative role in normal fibroblast physiology. FAP alpha inhibitors,

XX FT particularly a collagen derivative or (S)-valyl-pyrrolidine-2(R)-boronic

XX FT acid, may be used to treat various epithelial cancers, e.g. breast, lung,

XX FT skin, pancreas or colon cancer. Note: This sequence differs from the

XX FT amino acid sequence of FAP-alpha found in AAW31963 between residues 625

XX FT to 668

XX FT Sequence 760 AA;

Query Match 100.0%; Score 76; DB 2; Length 760;

Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEENT 15
| | | | | | | | | | | | | | | |
Db 24 IVLRPSRVHNSSEENT 38

RESULT 5

ID ABR47452 standard; protein; 760 AA.

XX AC ABR47452;

DT 12-JUN-2003 (first entry)

DE Breast cancer associated protein sequence SEQ ID NO:136.

XX Human; breast cancer; cytostatic; gene therapy.

XX OS Homo sapiens.

XX FN WO2003004989-A2.

XX PD 16-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019669.

XX PR 21-JUN-2001; 2001US-0299887P.

XX PR 27-JUN-2001; 2001US-0301572P.

XX PR 18-JUL-2001; 2001US-0306501P.

XX PR 25-SEP-2001; 2001US-0325002P.

XX PR 05-MAR-2002; 2002US-0362585P.

XX PR 14-MAY-2002; 2002US-0380391P.

XX FA (MILL-) MILLENIUM PHARM INC.

XX PI Lillie J, Gannavarapu M, Glatt K, Hoersht S, Kamatkar S;

XX PI Mertens M, Monahan JE, Myer Y, Xu Y, Zhao X, Meyers RE;

XX PI Bast RC, Horcobegyi GN, Fuszta L, Meric P, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

XX DR N-PSDB; ACC50144.

XX PT Breast cancer diagnosis or treatment by comparing the level of expression

XX PT of a marker in a patient sample with that in the control non-breast

XX PT cancer sample.

XX PS Claim 1; SEQ ID NO 136; 128pp; English.

XX CC The present invention describes a method for assessing whether a patient

XX CC is afflicted with breast cancer. The method comprises comparing the level

XX CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

XX CC ABR47386 to ABR47632) in a patient sample and the normal level of

XX CC expression of the marker in a control non-breast cancer sample, where a

XX CC significant increase in the level of expression of the marker in the

XX CC patient sample and the normal level is an indication that the patient is

XX CC afflicted with breast cancer. The breast cancer associated sequences from

XX CC the present invention have cytostatic activities and can be used in gene

XX CC therapy. The method is useful for diagnosing and treating breast cancer.

XX CC N.B. The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 760 AA;

Query Match 100.0%; Score 76; DB 6; Length 760;

Best Local Similarity 100.0%; Pred. No. 0.0001;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEENT 15
| | | | | | | | | | | | | | | |
Db 24 IVLRPSRVHNSSEENT 38

RESULT 6

ID ADN95552

XX ADN95552 standard; protein; 760 AA.

XX AC ADN95552;

XX DT 01-JUL-2004 (first entry)

XX DE Human BEC/LEC-related protein sequence SeqID475.

XX KW growth; differentiation; blood endothelial cell; BEC;

XX KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;

XX KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;

XX KW vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;

XX KW inflammatory disease; cancer metastasis; lymphatic system; human.

XX OS Homo sapiens.

XX FN WO2003080640-A1.

XX PD 02-OCT-2003.

XX PF 07-MAR-2003; 2003WO-US006900.

XX PR 07-MAR-2002; 2002US-0363019P.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PA (LICN) LICENTIA LTD.

XX PI Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;

XX WPI; 2003-876899/81.

XX DR N-PSDB; ADN95553.

XX PS Example 1; SEQ ID NO 475; 176pp; English.

This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprises contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGFR-3; and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotrophic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

XX SQ Sequence 760 AA;

Query Match 100.0%; Score 76; DB 7; Length 760;

Best Local Similarity 100.0%; Pred. No. 0.0001;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVLPRSRVHNSSENT 15
Db 24 IVLPRSRVHNSSENT 38

RESULT 7
ADQ21351
ID ADQ21351 standard; protein; 760 AA.
XX AC ADQ21351;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4171.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX OS Homo sapiens.
XX FN WO2004048938-A2.
XX PD 10-JUN-2004.
XX XX 26-NOV-2003; 2003WO-US038193.
XX PF 26-NOV-2002; 2002US-0429739P.
XX PR (PROT-) PROTEIN DESIGN LABS INC.
XX PA Aziz N, Ginsburg WM, Zlotnik A;
XX PI WPI; 2004-441208/41.
XX DR Early detection of soft tissue sarcoma comprises determining expression
XX PT of a gene in a first soft tissue sample and a normal soft tissue sample
XX FT and comparing the gene expression, also useful in treating soft tissue
XX FT sarcoma.

Example 2; SEQ ID NO 4171; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

Seq Sequence 760 AA;

Query Match 100.0%; Score 76; DB 8; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IVLPRSRVHNSSENT 15
Db 24 IVLPRSRVHNSSENT 38

RESULT 8
ADW14775
ID ADW14775 standard; protein; 760 AA.
XX AC ADW14775;
XX XX 07-APR-2005 (first entry)
XX DT

DE Tumor-associated antigenic target TAT430.
XX Tumor-associated antigenic target; TAT430; cancer; neoplasm; cytostatic;
KW ovary tumor; uterus tumor.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
FH 1..28
FT /label= Signal_peptide
FT 10..15
FT /note= "N-myristoylation site"
FT 29..760
FT /label= Mature_protein
FT 42..548
FT /note= "Dipeptidyl peptidase IV"
FT 49..52
FT /note= "N-glycosylation site"
FT 84..89
FT /note= "N-myristoylation site"
FT 92..95
FT /note= "N-glycosylation site"
FT 99..102
FT /note= "N-glycosylation site"
FT 159..164
FT /note= "N-myristoylation site"
FT 285..290
FT /note= "N-myristoylation site"
FT 314..317
FT /note= "N-glycosylation site"
FT 366..374
FT /note= "Tyrosine kinase phosphorylation site"
FT 402..410
FT /note= "Tyrosine kinase phosphorylation site"
FT 419..422
FT /note= "Amidation site"
FT 438..443
FT /note= "Cytochrome c family heme-binding site signature"
FT 470..475
FT /note= "N-myristoylation site"
FT 542..547
FT /note= "N-myristoylation site"
FT 552..629
FT /note= "Prolyl oligopeptidase family homology"
FT 566..571
FT /note= "N-myristoylation site"
FT 599..629
FT /note= "Prolyl endopeptidase family serine active site"
FT 626..631
FT /note= "N-myristoylation site"
FT 636..639
FT /note= "Glycosaminoglycan attachment site"
FT 639..644
FT /note= "N-myristoylation site"
FT 679..682
FT /note= "N-glycosylation site"
FT 738..743
FT /note= "N-myristoylation site"
XX WO2005003154-A2.
XX
XX 13-JAN-2005.
XX 02-JUL-2004; 2004WO-US021353.
XX 02-JUL-2003; 2003US-0484959P.
XX (SETH) GENENTECH INC.
XX Ashkenazi A, Cairns B, Dowd P, Frantz G, French D, Gonzalez L;
PI Polakls P, Smith V, Wolf B, Wu TD, Zhang Z;
XX WPI; 2005-091764/10.
XX DR

immune inhibition; fibroblast activation protein alpha dimer;
FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
immune disorder; neurological disease; autoimmune disease;
immunosuppressive; graft versus host disease; transplant rejection;
endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
cardiovascular disease; metabolic disorder; hashimoto's disease;
antithyroid; endocrine disease; inflammatory bowel disease;
antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
autoimmune hepatitis; antiinflammatory; hepatotropic;
systemic lupus erythematosus; dermatological; dermatological disease;
uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
hematological disease; rheumatic fever; antipyretic; Crohns disease;
psoriasis; antipsoriatic; graves disease; antithyroid;
respiratory syncytial virus infection; respiratory-gen.; virucide.

CC OS Homo sapiens.
XX XX WO2005071073-A1.
XX XX 04-AUG-2005.
XX XX 10-JAN-2005; 2005WO-US000709.
XX XX 09-JAN-2004; 2004US-0535577P.
XX XX (POIN-) POINT THERAPEUTICS INC.
XX XX Mclean PA, Jones B, Miller GT, Jesson MI;
XX XX WPI; 2005-564220/57.
XX XX Down-regulating an immune response comprises administering to a subject
XX XX in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
XX XX amount effective to down-regulate an immune response.
XX XX Claim 134; SEQ ID NO 61; 177pp; English.

CC The invention relates to a method of down-regulating an immune response,
CC which comprises administering to a subject a fibroblast activation
CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC an immune response. Also included are the following: a composition
CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC carrier, where the composition is sterile and lacks an adjuvant; a
CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC acceptable carrier, and a non-adjuvant second agent; a composition
CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
CC The method further comprises administering to the subject a second agent.
CC The second agent is an anti-inflammatory agent, immunosuppressant, or
CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
CC substitution relative to wild type FAP alpha dimer. The amino acid
CC substitution is present in the beta-propeller domain, the catalytic
CC domain, or an N-linked glycosylation site and alters disulfide bond
CC formation. The immune response is an especially an IL-1 mediated
CC condition, abnormal immune response selected from inflammation,
CC autoimmune disease, sepsis, graft versus host disease, transplant
CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
CC response is subsequent to an infection, such as an RSV infection. The
CC autoimmune disease is selected from c, autoimmune thyroiditis, systemic
CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
CC sclerosis. The subject does not have cancer or a predisposition to

CC cancer. The present sequence represents the amino acid sequence of human
CC soluble fibroblast activation protein (FAP) alpha dimer (corresponding to
CC AEB94159 minus N-terminal amino acids 1-26).
XX XX Sequence 734 AA;
SQ Query Match 84.2%; Score 64; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0;
QY 4 RPSRVHNSSEENT 15
Db 1 RPSRVHNSSEENT 12
|||||
RESULT 11
AEB94178
ID AEB94178 standard; protein; 50 AA.
XX XX AEB94178;
AC XX 06-OCT-2005 (first entry)
XX XX Mouse soluble FAP alpha dimer N-terminal fragment SEQ ID NO:21.
XX XX Immune inhibition; fibroblast activation protein alpha dimer;
XX FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
XX immune disorder; neurological disease; autoimmune disease;
XX immunosuppressive; graft versus host disease; transplant rejection;
XX endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
XX musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
XX inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
XX cardiovascular disease; metabolic disorder; hashimoto's disease;
XX antithyroid; endocrine disease; inflammatory bowel disease;
XX antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
XX rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
XX autoimmune hepatitis; antiinflammatory; hepatotropic;
XX systemic lupus erythematosus; dermatological; dermatological disease;
XX uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
XX hematological disease; rheumatic fever; antipyretic; Crohns disease;
XX psoriasis; antipsoriatic; graves disease; antithyroid;
XX respiratory syncytial virus infection; respiratory-gen.; virucide.
XX XX Mus musculus.
XX OS WO2005071073-A1.
XX XX 04-AUG-2005.
XX XX 10-JAN-2005; 2005WO-US000709.
XX XX 09-JAN-2004; 2004US-0535577P.
XX XX (POIN-) POINT THERAPEUTICS INC.
XX XX Mclean PA, Jones B, Miller GT, Jesson MI;
XX XX WPI; 2005-564220/57.
XX XX Down-regulating an immune response comprises administering to a subject
XX XX in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
XX XX amount effective to down-regulate an immune response.
XX XX Example 1.4; SEQ ID NO 21; 177pp; English.

CC The invention relates to a method of down-regulating an immune response,
CC which comprises administering to a subject a fibroblast activation
CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
CC an immune response. Also included are the following: a composition
CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
CC carrier, where the composition is sterile and lacks an adjuvant; a
CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
CC acceptable carrier, and a non-adjuvant second agent; a composition

comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising a FAP alpha dimer enzyme lacking amino acids 269-448 and comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, antiparasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPP1V/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the N-terminal 50 amino acids of mouse wild type FAP.

Sequence 50 AA;

Query Match 68.4%; Score 52; DB 9; Length 50;
Best Local Similarity 73.3%; Pred. No. 0.11;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSEENT 15
|||||||:
DB 24 IVLPRSRVYKPEGNT 38

RESULT 12
ID AEB94163
AC AEB94163
XX AEB94163;

06-OCT-2005 (first entry)

Mouse wild type fibroblast activation protein (FAP) alpha dimer.

immune inhibition; fibroblast activation protein alpha dimer;
FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
immune disorder; neurological disease; autoimmune disease;
immunosuppressive; graft versus host disease; transplant rejection;
endotoxic shock; osteoarthritis; antiallergic; osteopathic;
musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
inflammation; respiratory disease; atherosclerosis; arteriosclerotic;
cardiovascular disease; metabolic disorder; hashimoto's disease;
antithyroid; endocrine disease; inflammatory bowel disease;
antifibrotic; gastrointestinal-gen.; gastrointestinal disease;
rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
autoimmune hepatitis; antiinflammatory; hepatotropic;
systemic lupus erythematosus; dermatological; dermatological disease;
uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
hematological disease; rheumatic fever; antipyretic; Crohns disease;
psoriasis; antipsoriatic; graves disease; antithyroid;
respiratory syncytial virus infection; respiratory-gen.; virucide;
enzyme.

OS Mus musculus.

PN WO2005071073-A1.

XX 04-AUG-2005.

XX 10-JAN-2005; 2005WO-US000709.
XX 09-JAN-2004; 2004US-0535577P.
PR (POIN-) POINT THERAPEUTICS INC.
XX Mclean PA, Jones B, Miller GT, Jesson MI;
PI WPI; 2005-564220/57.
XX N-PSDB; AEB94162.

Down-regulating an immune response comprises administering to a subject in need a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response.

Disclosure; SEQ ID NO 6; 177pp; English.

The invention relates to a method of down-regulating an immune response, which comprises administering to a subject a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response. Also included are the following: a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, where the composition is sterile and lacks an adjuvant; a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, and a non-adjuvant second agent; a composition comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, antiparasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPP1V/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of mouse wild type fibroblast activation protein (FAP) alpha dimer.

Sequence 761 AA;

Query Match 68.4%; Score 52; DB 9; Length 761;
Best Local Similarity 73.3%; Pred. No. 2.4;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSEENT 15
|||||||:
DB 24 IVLPRSRVYKPEGNT 38

RESULT 13
ID AEB97649
XX AEB97649
XX AEB97649;

25-SEP-2001 (first entry)

XX

DE Human sperm protein 45.
XX Human sperm protein 45; recombinant production; malignant tumour; cancer;
KW blood disease; HIV infection; human immunodeficiency virus;
KW immune disorder; inflammatory condition; cytostatic; anti-HIV;
KW antiinflammatory; immunomodulator.
XX Homo sapiens.
OS WO200146236-A1.
XX 28-JUN-2001.
XX 18-DEC-2000; 2000WO-CN000583.
XX 22-DEC-1999; 99CN-00125689.
XX (BLOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX Mao Y, Xie Y;
XX WPI; 2001-441666/47.
XX N-PSDB; AAH46231.
XX Human sperm protein 45 and encoded polynucleotide, applicable in
PT diagnosis and treatment of cancer, hemopathy, HIV infection,
PT immunological diseases and various inflammation.
XX Claim 1; Page 32-33; 41pp; Chinese.
XX This sequence represents human sperm protein 45. The protein has a
CC molecular weight of 45 kD, and has 27% identity and 45% homology over a
CC 268 amino acid stretch with an Ensis minor male gonad nuclear protein
CC (GenBank accession number L41834). The invention relates to human sperm
CC protein 45 (AAB97649), nucleic acids encoding it (AAH46231), and a method
CC for the recombinant production of the protein. The present invention
CC additionally discloses an antagonist of sperm protein 45 for therapeutic
CC use, and an antibody which specifically binds to sperm protein 45. Sperm
CC protein 45, and nucleotides which encode it may be used for treating a
CC variety of diseases, such as malignant tumours, blood diseases, HIV
CC (human immunodeficiency virus) infection, immune disorders and
CC inflammatory conditions. The protein may also be used to screen for
CC modulators of its activity or for peptide fingerprinting identification.
CC The polynucleotide can be used as a primer for nucleic acid amplification
CC reactions or as a probe for hybridisation reactions, or in producing gene
CC chips or microarrays
XX Sequence 407 AA;
SQ Query Match 55.3%; Score 42; DB 4; Length 407;
Best Local Similarity 63.6%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPSRVHNSSEN 14
DB 279 RPSKLHNSKVN 289
RESULT 14
ADH13246
ID ADH13246 standard; protein; 425 AA.
XX ADH13246;
XX 11-MAR-2004 (first entry)
XX Human malignant neoplasia-related protein SeqID95.
DE malignant neoplasia; cytostatic; breast cancer; ovarian cancer;
KW gastric cancer; colon cancer; oesophageal cancer; mesenchymal cancer;
KW bladder cancer; non-small cell lung cancer; human.
XX Homo sapiens.
OS

XX EPL1365034-A2.
XX 26-NOV-2003.
XX 09-MAY-2003; 2003EP-00010447.
XX 21-MAY-2002; 2002EP-00010291.
XX 13-FEB-2003; 2003EP-00003112.
XX (FARB) BAYER AG.
XX Wirtz R, Munnes M, Kallabis H;
XX WPI; 2004-073279/08.
XX N-PSDB; ADH13223.
XX Predicting, diagnosing or prognosing malignant neoplasia by detecting at
PT least two markers, where the markers are genes from one or more
PT chromosomal regions altered in malignant neoplasia.
XX Claim 11; SEQ ID NO 95; 267pp; English.
XX This invention relates to a novel method for the prediction, diagnosis,
CC or prognosis of malignant neoplasia by the detection of at least two
CC markers. The invention may also be useful for the development of a
CC cytostatic compounds through the regulation of the expression of a gene
CC or activity of a protein associated with malignant neoplasia. The method
CC is useful for prediction, diagnosis or prognosis of malignant neoplasia,
CC such as breast cancer, ovarian cancer, gastric cancer, colon cancer,
CC oesophageal cancer, mesenchymal cancer, bladder cancer or non-small cell
CC lung cancer. The polynucleotides and polypeptides defined in the
CC specification, antisense polynucleotides targeting the polynucleotides,
CC antibodies targeting either one of the polynucleotides or polypeptides,
CC and compounds identified by the screening methods are useful for
CC preventing or treating malignant neoplasia. The disease treated is
CC preferably breast cancer. The present sequence is that of a human
CC malignant neoplasia-related protein which may be used in the method of
XX the invention.
SQ Sequence 425 AA;
Query Match 55.3%; Score 42; DB 8; Length 425;
Best Local Similarity 63.6%; Pred. No. 81;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPSRVHNSSEN 14
DB 297 RPSKLHNSKVN 307
RESULT 15
AEA15133
ID AEA15133 standard; protein; 425 AA.
XX AEA15133;
XX 28-JUL-2005 (first entry)
XX Human polypeptide #46.
XX Diagnosis; prognosis; cancer; breast tumor; ovary tumor; stomach tumor;
KW colon tumor; esophagus tumor; bladder tumor; non-small-cell lung cancer;
KW cytostatic; neoplasm.
XX Homo sapiens.
XX WO2005047534-A2.
XX 26-MAY-2005.
XX 15-OCT-2004; 2004WO-EP011599.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX Example 1; SEQ ID NO 6763; 1069pp; English.
XX Sequences AU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 119 AA;

Query Match 53.9%; Score 41; DB 4; Length 119;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHSEENT 15
Db 82 RPSRVHIDAAT 93
||||| : |

RESULT 18
ABM42087
ID ABM42087 standard; protein; 119 AA.
XX
AC ABM42087;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #6763.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliee-Douglas J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64455.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.
XX Example 1; SEQ ID NO 6763; 1481pp; English.
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 119 AA;

Query Match 53.9%; Score 41; DB 6; Length 119;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHSEENT 15
Db 82 RPSRVHIDAAT 93
||||| : |

RESULT 19
ADY08853
ID ADY08853 standard; protein; 458 AA.
XX
AC ADY08853;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 64668.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX

PN EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126783P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01324256P.
PR 14-MAY-1999; 99US-01324218P.
PR 14-MAY-1999; 99US-01324219P.
PR 14-MAY-1999; 99US-01324221P.
PR 14-MAY-1999; 99US-01324370P.
PR 18-MAY-1999; 99US-01324768P.
PR 19-MAY-1999; 99US-01324941P.
PR 20-MAY-1999; 99US-01325124P.
PR 21-MAY-1999; 99US-0132533P.
PR 24-MAY-1999; 99US-0132629P.
PR 25-MAY-1999; 99US-0132621P.
PR 27-MAY-1999; 99US-01326392P.
PR 28-MAY-1999; 99US-01326782P.
PR 01-JUN-1999; 99US-0132722P.
PR 03-JUN-1999; 99US-0132528P.
PR 04-JUN-1999; 99US-01327502P.
PR 07-JUN-1999; 99US-0132724P.
PR 08-JUN-1999; 99US-01328094P.
PR 10-JUN-1999; 99US-01328540P.
PR 10-JUN-1999; 99US-01328847P.
PR 14-JUN-1999; 99US-01329119P.
PR 16-JUN-1999; 99US-01329452P.
PR 16-JUN-1999; 99US-01329453P.
PR 17-JUN-1999; 99US-01329492P.
PR 18-JUN-1999; 99US-01329454P.
PR 18-JUN-1999; 99US-01329455P.
PR 18-JUN-1999; 99US-01329456P.
PR 18-JUN-1999; 99US-01329457P.
PR 18-JUN-1999; 99US-01329458P.
PR 18-JUN-1999; 99US-01329459P.
PR 18-JUN-1999; 99US-01329460P.
PR 18-JUN-1999; 99US-01329461P.
PR 18-JUN-1999; 99US-01329462P.
PR 18-JUN-1999; 99US-01329463P.
PR 18-JUN-1999; 99US-01329750P.
PR 18-JUN-1999; 99US-01329763P.
PR 21-JUN-1999; 99US-01329817P.
PR 21-JUN-1999; 99US-01329899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR	20-SEP-1999;	99US-0154779P.	PR	05-MAR-1999;	99US-0123180P.
PR	22-SEP-1999;	99US-0155133P.	PR	09-MAR-1999;	99US-0123548P.
PR	23-SEP-1999;	99US-0155486P.	PR	23-MAR-1999;	99US-0125788P.
PR	24-SEP-1999;	99US-0155659P.	PR	25-MAR-1999;	99US-0126264P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-MAR-1999;	99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999;	99US-0127462P.
PR	04-OCT-1999;	99US-0157117P.	PR	06-APR-1999;	99US-0128234P.
PR	05-OCT-1999;	99US-0157753P.	PR	08-APR-1999;	99US-0128714P.
PR	06-OCT-1999;	99US-0157855P.	PR	15-APR-1999;	99US-0129845P.
PR	07-OCT-1999;	99US-0158029P.	PR	16-APR-1999;	99US-0130077P.
PR	08-OCT-1999;	99US-0158232P.	PR	21-APR-1999;	99US-0130449P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999;	99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	23-APR-1999;	99US-0130891P.
PR	13-OCT-1999;	99US-0159294P.	PR	28-APR-1999;	99US-0131449P.
PR	13-OCT-1999;	99US-0159295P.	PR	30-APR-1999;	99US-0132048P.
PR	14-OCT-1999;	99US-0159329P.	PR	30-APR-1999;	99US-0132407P.
PR	14-OCT-1999;	99US-0159330P.	PR	04-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAY-1999;	99US-0132485P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999;	99US-0132486P.
PR	14-OCT-1999;	99US-0159638P.	PR	06-MAY-1999;	99US-0132487P.
PR	18-OCT-1999;	99US-0159584P.	PR	07-MAY-1999;	99US-0132863P.
PR	21-OCT-1999;	99US-0160741P.	PR	11-MAY-1999;	99US-0134256P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	25-OCT-1999;	99US-0161404P.	PR	24-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161406P.	PR	27-MAY-1999;	99US-0136392P.
PR	26-OCT-1999;	99US-0161359P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161360P.	PR	01-JUN-1999;	99US-0137222P.
PR	26-OCT-1999;	99US-0161361P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161922P.	PR	07-JUN-1999;	99US-0137724P.
PR	28-OCT-1999;	99US-0161993P.	PR	08-JUN-1999;	99US-0138094P.
PR	29-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138540P.
Query Match 53.9%; Score 41; DB 3; Length 835;					99US-0138847P.
Best Local Similarity 60.0%; Pred. No. 2.7e+02;					99US-0139119P.
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;					99US-0139199P.
Qy	5 PSRVHNSSEN 14				99US-0139452P.
Db	564 PSKIHNSSEN 573				99US-0139453P.
RESULT 24					99US-0139492P.
AAG31411					99US-0139454P.
ID	AAG31411 standard; protein; 929 AA.				99US-0139455P.
XX	AC				99US-0139456P.
XX	AAG31411;				99US-0139457P.
XX	17-OCT-2000 (first entry)				99US-0139458P.
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 37716.				99US-0139460P.
DE	Arabidopsis thaliana				99US-0139461P.
XX	Protein identification; signal transduction pathway; metabolic pathway;				99US-0139462P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				99US-0139463P.
KW	termination sequence.				99US-0139454P.
XX	Arabidopsis thaliana.				99US-0139455P.
OS	EP1033405-A2.				99US-0139457P.
XX	06-SEP-2000.				99US-0139458P.
PN	25-FEB-2000; 2000EP-00301439.				99US-0139461P.
PD	99US-0140353P.				99US-0139462P.
XX	99US-0140354P.				99US-0139463P.
XX	99US-0140695P.				99US-0139750P.
XX	99US-0140823P.				99US-0139763P.
OS	99US-0140991P.				99US-0139817P.
XX	99US-0141287P.				99US-0139899P.
XX	99US-0141842P.				99US-0140353P.
PN	99US-0142154P.				99US-0140354P.
XX	99US-0142055P.				99US-0140695P.
PD	99US-0142390P.				99US-0140823P.
XX	99US-0142803P.				99US-0140991P.
PF	99US-0142920P.				99US-0141287P.
XX	99US-0142977P.				99US-0141842P.
PR	99US-0143542P.				99US-0142154P.

PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.

PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 53.9%; Score 41; DB 3; Length 929;

Best Local Similarity 60.0%; Fred. No. 3e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 6; Conservative 2;

QY 5 PSRVHNSSEN 14

||:||||

Db 658 PSKIHNSSEN 667

RESULT 25

AAG31410

ID AAG31410 standard; protein; 937 AA.

XX AAG31410;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 37715.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 37715.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	99US-0132487P.	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134218P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134219P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	99US-0134768P.	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	99US-0134941P.	PR	02-AUG-1999;	99US-0146386P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146389P.
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	99US-0136382P.	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	99US-0137724P.	PR	08-AUG-1999;	99US-0147493P.
PR	08-JUN-1999;	99US-0138094P.	PR	08-AUG-1999;	99US-0147335P.
PR	10-JUN-1999;	99US-0138540P.	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	99US-0139492P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139454P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139455P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139456P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139459P.	PR	23-AUG-1999;	99US-0149802P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139461P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139462P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	05-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	06-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144003P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144085P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144332P.			

XX Ubiquitin-specific C terminal hydrolyzing protease gene regulating pseudo
PT -arabis protein degradation.
PT
XX
XX Claim 2; Page 7 (Disclosure); 10pp; Chinese.
PS
XX

Query Match 53.9%; Score 41; DB 5; Length 2828;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 7; Mismatches 2; Indels

```

Qy      1 IVLRPSRVHNSSEN 14
      :|::|::|::|::|
Db      2238 VVMPKAKIEHKEEN 2251

RESULT 28
ABG31319
ID      ABG31319 standard; protein; 2828 AA.
XX
XX
AC      ABG31319;
XX
XX      05-NOV-2002 (first entry)
Dt      Human adlcan protein.
XX
XX      Human; mechanical stress; protein 608; OCP; CMF608; osteoporosis;
Kw      bone fracture; bone elongation; osteopenia; osteoarthritis;
Kw      osteopetrosis; osteosclerosis; periodontitis; low bone density;
Kw      osteopathic; antiinflammatory; antiarthritic; adlcan.
XX
XX      Homo sapiens.
XX
XX      WO200246364-A2.
Pn
XX
XX      13-JUN-2002.
Pd
XX
XX      04-DEC-2001; 2001WO-US046400.
Pf
XX
XX      04-DEC-2000; 2000US-00729485.
Pr      08-MAR-2001; 2001US-00802318.
Pr      13-JUL-2001; 2001US-00905129.
Pr      06-NOV-2001; 2001US-00991630.
XX
XX      (QUAR-) QUARK BIOTECH INC.
Pa
XX
XX      Einat P, Segev O, Skaliter R, Feinstein E, Faerman A, Samach A;
Pi
XX
XX      WPI; 2002-590536/63.
Dr      N-PSDB; ABK90044.
XX
XX      New isolated mechanical stress induced polynucleotide, useful for
Pt      diagnosis, treatment, prevention or control of osteoporosis, osteopenia,
Pt      osteopetrosis, osteosclerosis, osteoarthritis, periodontitis, bone
Pt      fracture.
XX
XX      Claim 33; Fig 51; 164pp; English.
Ps
XX
XX      The present invention relates to the isolation of genes associated with
Cc      mechanical stress, and the proteins encoded by them. The genes of the
Cc      invention are designated protein 608, OCP or CMF608. The OCP
Cc      polynucleotide and polypeptide sequences are useful for preventing,
Cc      treating or controlling osteoporosis, bone fractures, bone elongation,
Cc      osteopenia, osteoarthritis, osteopetrosis, osteosclerosis, periodontitis,
Cc      low bone density or other conditions involving mechanical stress in a
Cc      subject. The present sequence represents human adlcan protein
XX
XX      Sequence 2828 AA;
Sq
      Query Match      53.9%; Score 41; DB 5; Length 2828;
      Best Local Similarity 35.7%; Pred. No. 1.1e+03;
      Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IVLRPSRVHNSSEN 14
      :|::|::|::|
Db      2238 VVMPKAKIEHKEEN 2251

RESULT 29
ABG32897
ID      ABG32897 standard; protein; 2828 AA.
XX
XX
AC      ABG32897;

```

```

XX
Dt      27-NOV-2002 (first entry)
XX
XX      Human Adlcan protein.
XX
XX      Human; OCP; osteoclast protein; osteoporosis; osteopenia; osteopetrosis;
Kw      osteosclerosis; osteoarthritis; periodontosis; bone fracture;
Kw      mechanical stress; adlcan; gene therapy.
XX
XX      Homo sapiens.
Os
XX
XX      US2002086825-A1.
Pn
XX
XX      04-JUL-2002.
Pd
XX
XX      08-MAR-2001; 2001US-00802318.
Pf
XX
XX      11-MAY-1998; 98US-0084944P.
Pr      15-MAY-1998; 98US-0085673P.
Pr      30-MAY-2000; 2000US-0207821P.
Pr      04-DEC-2000; 2000US-00729485.
XX
XX      (EINA/) EINAT P.
Pa      (SEGE/) SEGEV O.
Pa      (SKAL/) SKALITER R.
Pa      (FEIN/) FEINSTEIN E.
Pa      (FAER/) FAERMAN A.
XX
XX      Einat P, Segev O, Skaliter R, Feinstein E, Faerman A;
Pi
XX
XX      WPI; 2002-681611/73.
Dr      N-PSDB; ABS54195.
XX
XX      Novel isolated mechanical stress induced polypeptide useful for
Pt      preventing, treating or controlling osteoporosis, bone elongation,
Pt      osteopenia, periodontitis, low bone density or for fracture healing.
XX
XX      Disclosure; Fig 51; 198pp; English.
XX
XX      The invention relates to a composition comprising an isolated mechanical
Cc      stress induced polypeptide identified as adlcan or its functional
Cc      portion or a polypeptide which is at least homologous or identical to it,
Cc      and a physiologically acceptable excipient. Also included are an anti-
Cc      adlcan antibody or composition containing it, a composition comprising
Cc      an isolated nucleic acid molecule encoding adlcan and a physiologically
Cc      acceptable excipient and a composition comprising a vector that comprises
Cc      the adlcan nucleic acid and a physiologically acceptable excipient. The
Cc      compositions and antibody are useful for preventing, treating or
Cc      controlling osteoporosis, or for fracture healing, bone elongation,
Cc      osteopenia, periodontosis, low bone density, osteopetrosis,
Cc      osteosclerosis, osteoarthritis or other conditions involving mechanical
Cc      stress or its lack in a subject. The nucleic acid is useful in diagnosis
Cc      (for instance risk determination), treatment, prevention, or control, of
Cc      the above diseases. Proteins homologous to adlcan include OCP
Cc      (osteoclast protein). The present sequence is the human adlcan protein
XX
XX      Sequence 2828 AA;
Sq
      Query Match      53.9%; Score 41; DB 5; Length 2828;
      Best Local Similarity 35.7%; Pred. No. 1.1e+03;
      Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy      1 IVLRPSRVHNSSEN 14
      :|::|::|::|
Db      2238 VVMPKAKIEHKEEN 2251

RESULT 30
AD117071
ID      AD117071 standard; protein; 2828 AA.
XX
XX
AC      AD117071;

```

DT	15-APR-2004	(first entry)	PA	(CURA-) CURAGEN CORP.
XX	Human NOVX protein homologue SeqID 607.		XX	Tchernev VT, Spytek KA, Zethusen BD, Patturajan M, Shimkets RA;
DE	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;		PI	Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
XX	inflammation; autoimmune disorder; allergy; blood disorder;		PI	Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;		PI	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;		XX	WPI; 2002-706998/76.
KW	Alzheimer's disease; infection; str.		DR	
XX	Homo sapiens.		XX	New NOVX polypeptides and nucleic acids, useful for preventing or
XX	WO200268649-A2.		PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX	06-SEP-2002.		PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX			PT	pharmacogenomics.
XX			PS	Disclosure; SEQ ID NO 607; 1498pp; English.
XX	31-JAN-2002; 2002WO-US002785.		XX	This invention relates to a novel nucleic acids, and encoded polypeptides
XX	31-JAN-2001; 2001US-0265395P.		CC	thereof, which have properties related to the stimulation of biochemical
PR	31-JAN-2001; 2001US-0265412P.		CC	or physiological responses in a cell, tissue, organ or organism.
PR	31-JAN-2001; 2001US-0265514P.		CC	Specifically, it refers to the use of biologically active fragments for
PR	31-JAN-2001; 2001US-0265517P.		CC	diagnostic and prognostic assays and furthermore in the treatment of
PR	02-FEB-2001; 2001US-0266408P.		CC	diverse pathological conditions. The present invention describes novel
PR	05-FEB-2001; 2001US-0266767P.		CC	human and murine NOVX proteins, as well as methods to modulate their
PR	07-FEB-2001; 2001US-0266975P.		CC	expression using antisense oligos, ribozymes and peptide nucleic acids.
PR	07-FEB-2001; 2001US-0267057P.		CC	The NOVX polypeptides, polynucleotides and antibodies are useful in
PR	08-FEB-2001; 2001US-0267459P.		CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
PR	09-FEB-2001; 2001US-0267823P.		CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in
PR	15-FEB-2001; 2001US-0268974P.		CC	treating or preventing diseases such as inflammation, autoimmune
PR	26-FEB-2001; 2001US-0271664P.		CC	disorders, allergies, blood disorders, acquired immunodeficiency syndrome
PR	27-FEB-2001; 2001US-0271833P.		CC	(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
PR	02-MAR-2001; 2001US-0272788P.		CC	arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
PR	02-MAR-2001; 2001US-0273048P.		CC	and epilepsy. Accordingly, these molecules have many activities including
PR	14-MAR-2001; 2001US-0275925P.		CC	cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
PR	14-MAR-2001; 2001US-0275947P.		CC	haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
PR	14-MAR-2001; 2001US-0275989P.		CC	antiaesthetic, nephrotropic, antiarthritic, hepatotropic.
PR	15-MAR-2001; 2001US-0276448P.		CC	neuroprotective, nootropic, antibacterial, virucide, antiparasitic.
PR	15-MAR-2001; 2001US-0276450P.		CC	relaxant and anticonvulsant. In addition, they are useful in screening
PR	16-MAR-2001; 2001US-0276397P.		CC	assays to identify small molecules that modulate or inhibit, for example,
PR	16-MAR-2001; 2001US-0276768P.		CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also
PR	20-MAR-2001; 2001US-0278652P.		CC	used as in chromosome mapping, tissue typing, preventive medicine and
PR	26-MAR-2001; 2001US-0278752P.		CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
PR	26-MAR-2001; 2001US-0278778P.		CC	of the invention.
PR	29-MAR-2001; 2001US-0279882P.		XX	Sequence 2828 AA;
PR	30-MAR-2001; 2001US-0280147P.		SQ	
PR	11-APR-2001; 2001US-0282992P.			Query Match 53.9%; Score 41; DB 5; Length 2828;
PR	11-APR-2001; 2001US-0283083P.			Best Local Similarity 35.7%; Pred. No. 1.1e+03;
PR	20-APR-2001; 2001US-0285133P.			Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
PR	23-APR-2001; 2001US-0285749P.			
PR	03-MAY-2001; 2001US-0288327P.			OY 1 IVLRPSRVHSEEN 14
PR	29-MAY-2001; 2001US-0288504P.			: : : : :
PR	30-MAY-2001; 2001US-0294473P.			Db 2238 VVMKPAKIEHKEEN 2251
PR	08-JUN-2001; 2001US-0296964P.			: : : : :
PR	18-JUN-2001; 2001US-0298959P.			
PR	19-JUN-2001; 2001US-0299324P.			RESULT 31
PR	13-AUG-2001; 2001US-0312020P.			AD117072
PR	16-AUG-2001; 2001US-0312889P.			ID AD117072 standard; protein; 2828 AA.
PR	16-AUG-2001; 2001US-0312908P.			XX AD117072;
PR	21-AUG-2001; 2001US-0313390P.			XX
PR	28-AUG-2001; 2001US-0315470P.			DT 15-APR-2004 (first entry)
PR	31-AUG-2001; 2001US-0316447P.			XX Human NOVX protein homologue SeqID 608.
PR	07-SEP-2001; 2001US-0318115P.			XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
PR	07-SEP-2001; 2001US-0318118P.			KW inflammation; autoimmune disorder; allergy; blood disorder;
PR	12-SEP-2001; 2001US-0318740P.			KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
PR	19-SEP-2001; 2001US-0323379P.			KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
PR	18-OCT-2001; 2001US-0330245P.			XX Alzheimer's disease; infection; str.
PR	14-NOV-2001; 2001US-0332701P.			OS Homo sapiens.
XX				XX
XX				PN WO200268649-A2.

XX PS Disclosure; Page 149-155; 233pp; English.

XX CC The invention comprises a method for assessing whether a patient is

XX CC afflicted with breast cancer or ovarian cancer. The method involves the

XX CC use of specific DNA markers. The method of the invention is useful in the

XX CC detection and treatment of ovarian and breast cancer. Amino acid

XX CC sequences ABJ37025 - ABJ37080 represent human breast/ovarian cancer-

XX CC related proteins.

XX SQ Sequence 2828 AA;

Query Match 53.9%; Score 41; DB 6; Length 2828;

Best Local Similarity 35.7%; Pred. No. 1.1e+03;

Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSEN 14

Db 2238 VVMKPAKIEHKEEN 2251

RESULT 33

ABU56609 standard; protein; 2828 AA.

XX AC ABU56609;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polypeptide #202.

XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;

XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334370P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX DR N-PSDB; ABX76338.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient

XX PT for treating lung cancer, by contacting a biological sample from the

XX PT patient with a polynucleotide that exhibits increased or decreased

XX PT expression in lung cancer.

XX PS Claim 27; Page 340-341; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated

XX CC transcript in a cell from a patient, comprising contacting a biological

XX CC sample from the patient with a polynucleotide that selectively hybridises

XX CC to a sequence that is at least 80 % identical to a gene that exhibits

XX CC increased or decreased expression in lung cancer samples. Lung cancer-

XX CC associated polynucleotides and polypeptides are used for identifying a

XX CC compound that modulates a lung cancer-associated polypeptide, for

XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung

XX CC cancer in a patient and for treating a mammal having lung cancer by

XX CC administering a modulatory compound identified. The methods are useful

XX CC for treating lung cancer, such as small cell lung cancer, non-small cell

XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,

XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,

XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and

XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

XX CC for diagnostic purposes and as targets for screening for therapeutic

XX CC compounds that modulate lung cancer, such as antibodies. Sequences

XX CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the

XX CC invention

XX SQ Sequence 2828 AA;

Query Match 53.9%; Score 41; DB 6; Length 2828;

Best Local Similarity 35.7%; Pred. No. 1.1e+03;

Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSEN 14

Db 2238 VVMKPAKIEHKEEN 2251

RESULT 34

ABR47439 standard; protein; 2828 AA.

XX AC ABR47439;

XX DT 12-JUN-2003 (first entry)

XX DE Breast cancer associated protein sequence SEQ ID NO:110.

XX KW Human; breast cancer; cytostatic; gene therapy.

XX OS Homo sapiens.

XX PN WO2003004989-A2.

XX PD 16-JAN-2003.

XX PF 21-JUN-2002; 2002WO-US019669.

XX PR 21-JUN-2001; 2001US-0299887P.

XX PR 27-JUN-2001; 2001US-0301572P.

XX PR 18-JUL-2001; 2001US-0306501P.

XX PR 25-SEP-2001; 2001US-0325002P.

XX PR 05-MAR-2002; 2002US-0362585P.

XX PR 14-MAY-2002; 2002US-0380391P.

XX PA (MILL-) MILLENIUM PHARM INC.

XX PI Lillie J, Gannavarapu M, Glatt K, Hoerish S, Kamatkar S;

XX PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

XX PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

XX DR N-PSDB; ACC50131.

XX PT Breast cancer diagnosis or treatment by comparing the level of expression

XX PT of a marker in a patient sample with that in the control non-breast

XX PT cancer sample.

XX PS Claim 1; SEQ ID NO 110; 128pp; English.

XX CC The present invention describes a method for assessing whether a patient

XX CC is afflicted with breast cancer. The method comprises comparing the level

XX CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

XX CC ABR47386 to ABR47632) in a patient sample and the normal level of

XX CC expression of the marker in a control non-breast cancer sample, where a

XX CC significant increase in the level of expression of the marker in the

XX CC patient sample and the normal level is an indication that the patient is

fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
inflammatory disease; autoimmune disease;
retinal neovascularisation syndrome; scarring; uterine fibroid;
detection; diagnosis; prognosis; drug screening; drug targeting;
wound healing; contraception; cytostatic; cardiant; immunomodulatory;
vulnery; gene therapy; vaccine.

Homo sapiens.

WO2003042661-A2.

22-MAY-2003.

13-NOV-2002; 2002WO-US036810.

13-NOV-2001; 2001US-0350666P.

21-NOV-2001; 2001US-0332464P.

29-NOV-2001; 2001US-0334393P.

03-DEC-2001; 2001US-0335394P.

14-DEC-2001; 2001US-0340376P.

08-JAN-2002; 2002US-0347211P.

10-JAN-2002; 2002US-0347349P.

08-FEB-2002; 2002US-035250P.

13-FEB-2002; 2002US-0356714P.

20-FEB-2002; 2002US-0359077P.

29-MAR-2002; 2002US-0368809P.

04-APR-2002; 2002US-0370110P.

12-APR-2002; 2002US-0372246P.

05-JUN-2002; 2002US-0386614P.

16-JUL-2002; 2002US-0396839P.

22-JUL-2002; 2002US-0397775P.

22-JUL-2002; 2002US-0397845P.

09-SEP-2002; 2002US-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

WPI; 2003-468649/44.

N-PSDB; ADN39639.

Determining the presence or absence of a pathological cell in a patient,
useful for diagnosing, prognosing or treating cancer, comprises detecting
a nucleic acid in a biological sample.

Claim 12; SEQ ID NO C227; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
whose expression is upregulated or downregulated in specific cancers or
other diseases such as angiogenic or fibrotic disorders, and to methods
of determining the presence or absence of a pathological cell in a
patient by detecting a nucleic acid at least 80% identical to those of
the invention or by detecting a polypeptide of the invention. The
invention also relates to expression vectors and host cells comprising a
nucleic acid of the invention; antibodies which specifically bind a
polypeptide of the invention; use of such antibodies for drug targeting;
and methods of screening for modulators of activity or expression of the
invention.

Query Match 53.9%; Score 41; DB 7; Length 2828;

Best Local Similarity 35.7%; Pred. No. 1.1e-03;

Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

1 IVLRPSRVHNSSEN 14

Db 2238 VMPAKIEHKEEN 2251
RESULT 37
ADN38740
ID ADN38740 standard; protein; 2828 AA.
XX
AC ADN38740;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:58.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
WPI; 2003-468649/44.
DR N-PSDB; ADN38739.
XX
PT Determining the presence or absence of a pathological cell in a patient,
useful for diagnosing, prognosing or treating cancer, comprises detecting
a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO 58; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
whose expression is upregulated or downregulated in specific cancers or
other diseases such as angiogenic or fibrotic disorders, and to methods
of determining the presence or absence of a pathological cell in a
patient by detecting a nucleic acid at least 80% identical to those of
the invention or by detecting a polypeptide of the invention. The
invention also relates to expression vectors and host cells comprising a
nucleic acid of the invention; antibodies which specifically bind a
polypeptide of the invention; use of such antibodies for drug targeting;
and methods of screening for modulators of activity or expression of the
invention.

XX Detecting expression of one or more nucleic acid sequences in biological
PT sample, useful for detecting cancer, comprises detecting a change in the
PT expression level of one or more nucleic acid sequences relative to a
PT control expression level.
XX
PS Claim 20; SEQ ID NO 124; 256pp; English.
XX
CC The invention relates to detecting differential expression of one or more
CC nucleic acid sequences (appearing as ABA04356-AEA04448 in a biological
CC sample comprising obtaining the sample from a subject, and detecting a
CC change in the expression level of one or more nucleic acid sequences
CC relative to a control expression level of one or more nucleic acid sequences, is
CC new. Also included are detecting cancer (or a pre-malignant condition
CC thereof) in a subject (comprising comparing the expression level of one
CC or more nucleic acid sequences in a biological sample from the subject
CC with a control expression level of the nucleic acid sequences, where a
CC change of at least two-fold in the expression level of the nucleic acid
CC sequences is indicative of cancer or pre-malignant condition), monitoring
CC the onset (or progression, or regression) of cancer (or a pre-malignant
CC condition) in a subject (by detecting in a biological sample of the
CC subject at a first point in time the expression of one or more nucleic
CC acid sequences, repeating the first step at a subsequent point in time
CC and comparing the expression level detected, where a change in the
CC expression level is indicative of progression of cancer or its pre-
CC malignant condition in the subject), determining prognosis for cancer or
CC its pre-malignant condition in a subject (comprising detecting in a
CC biological sample of the subject, the expression level of one or more
CC nucleic acid sequences, comparing the expression level with a reference
CC expression level of the nucleic acid sequences and evaluating the
CC prognosis of the subject based on the comparison), determining the
CC efficacy of a test compound for inhibiting cancer in a subject,
CC determining the efficacy of a therapy for inhibiting cancer in a subject,
CC selecting a composition for inhibiting cancer in a subject, inhibiting
CC cancer in a subject, a polypeptide encoded by the nucleic acids above
CC (appearing as ABA04449-AEA04541), an antibody that specifically binds to
CC the polypeptide sequence, and detecting in a biological sample the
CC presence of a polypeptide. The method is useful for detecting
CC differential expression of one or more nucleic acid sequences in a
CC biological sample, which is useful for detecting cancer (especially colon
CC cancer), monitoring the onset, progression, or regression of cancer or a
CC pre-malignant condition, or determining prognosis for cancer or its pre-
CC malignant condition in a subject, or for determining the efficacy of a
CC test compound for inhibiting cancer in a subject. The compound is useful
CC for inhibiting cancer in a subject. The antibodies may also be used to
CC treat cancer. The present sequence is a protein from a human gene over-
CC expressed in cancer samples.
XX
SQ Sequence 2828 AA;
Query Match 53.9%; Score 41; DB 9; Length 2828;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 IVLRPSRVHNSSEN 14
:|:|:|:|:|:|
DB 2238 VMKPAKIEHKEEN 2251
RESULT 43
ABR40110
ID ABR40110 standard; protein; 2853 AA.
XX
AC ABR40110;
XX
XX 04-JUL-2003 (first entry)
XX
DE Human cell adhesion and extracellular matrix protein, CADECM-7.
XX
DE Human; anti-HIV; anti-allergic; cerebroprotective; antiparkinsonian;
KW anticonvulsant; nootropic; neuroprotective; immunosuppressive;
KW dermatological; anti-inflammatory; cytostatic; antiarteriosclerotic;
KW gene therapy; cell adhesion; extracellular matrix; CADECM;
KW

KW Immune system disorder; AIDS; allergy; neurological disorder; stroke;
KW Parkinson's disease; epilepsy; developmental disorder; Down's syndrome;
KW cerebral palsy; connective tissue disorder; systemic lupus erythematosus;
KW genetic disorder; Alport's syndrome; cell proliferative disorder; cancer;
KW atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO2003027230-A2.
XX
PD 03-APR-2003.
XX
PF 02-AUG-2002; 2002WO-US024649.
XX
XX 03-AUG-2001; 2001US-0309964P.
PR 03-AUG-2001; 2001US-0310119P.
PR 17-AUG-2001; 2001US-0313091P.
PR 31-AUG-2001; 2001US-0316771P.
PR 07-SEP-2001; 2001US-0317896P.
PR 21-SEP-2001; 2001US-0324781P.
PR 05-OCT-2001; 2001US-0327606P.
PR 12-OCT-2001; 2001US-0328960P.
PR 09-NOV-2001; 2001US-034471P.
PR 17-MAY-2002; 2002US-0381291P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Burford N, Warren BA, Duggan BM, Mason PM, Richardson TW, Yue H;
PI Forsythe J, Elliott VS, Griffin JA, Gorvad AB, Azimzai Y;
PI Kalllick DA, Xu Y, Honchell CD, Baughn MR, Gietzen KJ, Lee S;
PI Walia NK, Tang YT, Nguyen DB, Becha SD, Lee SY, Ramkumar J;
XX
DR WPI; 2003-354645/33.
XX N-PSDB; ACC00398.
XX
XX New human cell adhesion and extracellular matrix proteins (CADECM),
PT useful for diagnosing, treating or preventing disorders associated with
PT aberrant CADECM expression, e.g. cancer, AIDS, atherosclerosis, allergies
PT or stroke.
XX
XX Claim 1; Page 178-185; 234pp; English.
XX
XX The present invention relates to novel human cell adhesion and
CC extracellular matrix proteins, CADECM-1 to CADECM-22, and their coding
CC sequences (ACC00392-ACC00413 and ABR40104-ABR40125). The coding sequences
CC and proteins are useful in diagnosing, treating and preventing disorders
CC associated with aberrant expression of CADECM, such as immune system
CC disorders (e.g. AIDS or allergies), neurological disorders (e.g. Down's
CC Parkinson's disease or epilepsy), developmental disorders (e.g. stroke,
CC syndrome or cerebral palsy), connective tissue disorders (e.g. systemic
CC lupus erythematosus), genetic disorders (e.g. Alport's syndrome) or cell
CC proliferative disorders (e.g. cancer or atherosclerosis)
XX
SQ Sequence 2853 AA;
Query Match 53.9%; Score 41; DB 6; Length 2853;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 IVLRPSRVHNSSEN 14
:|:|:|:|:|:|
DB 2263 VMKPAKIEHKEEN 2276
RESULT 44
ABG23884
ID ABG23884 standard; protein; 2993 AA.
XX
AC ABG23884;
XX
XX 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #23875.

PA (TOOL-) TOOLGEN INC.
XX Kim J, Lee D, Park J, Kim Y;
PI WPI; 2004-248440/23.
XX
PT Evaluating cellular genes by expressing nucleic acid encoding different
PT artificial, chimeric transcription factor in cells and identifying set of
PT genes altered by transcription factors.
XX
XX Disclosure; SEQ ID NO 7; 124pp; English.
XX
XX The invention relates to evaluating several cellular genes. The method
CC involves providing several cells containing nucleic acid encoding
CC different artificial, chimeric transcription factor, expressing nucleic
CC acids in cells, evaluating expression of several genes in cells to
CC provide expression information and identifying from expression
CC information, a set of two or more genes whose expression is altered by
CC first and second transcription factors. The nucleic acids encoding the
CC different artificial, chimeric transcription factors comprise a
CC preselected set of nucleic acids, where each member of the preselected
CC set of nucleic acids encodes an artificial, chimeric transcription factor
CC that can alter a cellular phenotype. Each artificial, chimeric
CC transcription factor comprises a first zinc finger domain, and further
CC comprises a second and third zinc finger domains. The first zinc finger
CC domain is a naturally occurring zinc finger domain. The first zinc finger
CC domain is a mammalian zinc finger domain. Each cell of several cells is
CC an animal or human cell. The method is useful for evaluating several
CC cellular genes. The present sequence represents the amino acid sequence
CC of a repression domain fragment of the UME6 protein.
XX
XX Sequence 90 AA;
Query Match 52.6%; Score 40; DB 8; Length 90;
Best Local Similarity 64.3%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
QY 2 VLRP--SRVHNSSE 13
DB ||||| :||||:
64 VLRPILLRIHNSSEQ 77
RESULT 49
ADP71011
ID ADP71011 standard; protein; 90 AA.
XX
XX AC ADP71011;
XX
XX DT 09-SEP-2004 (first entry)
XX
XX DE Novel zinc finger protein-related yeast UME6 repression domain SeqID119.
XX
XX KW DNA binding domain; zinc finger domain; VEGF; VEGF-A;
KW transcription activation domain; transcription repression domain;
KW protein transduction domain; p65; VP16 activation domain; Kid;
KW KOX repression domain; TAR protein; VP22 protein;
KW Antennapedia homeodomain; cytosolic; antiinflammatory; angiogenesis;
KW neoplastic disorder; inflammatory disorder; angiogenesis-based disorder.
XX
XX OS Saccharomyces cerevisiae.
XX
XX PN WO2004053130-A1.
XX
XX PD 24-JUN-2004.
XX
XX XX 09-DEC-2003; 2003WO-KR002693.
XX
XX PR 09-DEC-2002; 2002US-0431892P.
XX
XX PA (TOOL-) TOOLGEN INC.
XX
XX PI Kim J, Shin H, Kwon H;
XX
XX WPI; 2005-031673/03.
XX
XX New chimeric protein comprising zinc finger domains and a heterologous
PT protein transduction domain, useful in preparing a composition for
PT treating a subject having or being suspected of having neoplastic or
PT inflammatory disorder.

DR WPI; 2004-468871/44.
XX
XX New polypeptide comprising a DNA binding domain that includes a plurality
PT of zinc finger domains, and at least two of the zinc finger domains each
PT include respective zinc finger domain motifs, useful for modulating
PT angiogenesis.
XX
XX Disclosure; SEQ ID NO 119; 176pp; English.
XX
XX This invention relates to a novel polypeptide comprising a DNA binding
CC domain that includes a plurality of zinc finger domains, where the DNA
CC binding domain can bind to a site in a VEGF gene, and at least two of the
CC zinc finger domains each include respective zinc finger domain motifs
CC listed in the specification. The VEGF gene is the human VEGF-A gene. The
CC polypeptide regulates the VEGF gene expression. The polypeptide further
CC comprises a transcription activation domain, a transcription repression
CC domain, or a protein transduction domain. The transcription activation
CC domain comprises p65 or VP16 activation domain. The transcription
CC repression domain comprises Kid or KOX repression domain. The protein
CC transduction domain is a part of TAR protein, VP22 protein, or
CC Antennapedia homeodomain. The invention may be useful for the production
CC of compounds with a cytostatic or antiinflammatory activity acting as
CC angiogenesis modulators. The polypeptides and nucleic acids are useful
CC for modulating angiogenesis whilst the compounds may be useful for
CC preventing or treating a neoplastic disorder, an inflammatory disorder,
CC or an angiogenesis-based disorder. The present sequence is that of a
CC yeast UME6 protein repression domain which is related to the invention.
XX
XX Sequence 90 AA;
Query Match 52.6%; Score 40; DB 8; Length 90;
Best Local Similarity 64.3%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
QY 2 VLRP--SRVHNSSE 13
DB ||||| :||||:
64 VLRPILLRIHNSSEQ 77
RESULT 50
ADV39071
ID ADV39071 standard; protein; 90 AA.
XX
XX AC ADV39071;
XX
XX DT 10-MAR-2005 (first entry)
XX
XX DE Yeast UME6 repression domain.
XX
XX KW Cellular transport; neoplasm; inflammation; angiogenesis disorder; UME6;
KW cytosolic; antiinflammatory; antiangiogenic.
XX
XX OS Saccharomyces cerevisiae.
XX
XX PN WO2004108883-A2.
XX
XX PD 16-DEC-2004.
XX
XX PF 10-JUN-2004; 2004WO-KR001385.
XX
XX PR 10-JUN-2003; 2003US-0477459P.
XX
XX PA (TOOL-) TOOLGEN INC.
XX
XX PI Kim J, Shin H, Kwon H;
XX
XX WPI; 2005-031673/03.
XX
XX New chimeric protein comprising zinc finger domains and a heterologous
PT protein transduction domain, useful in preparing a composition for
PT treating a subject having or being suspected of having neoplastic or
PT inflammatory disorder.

PS Disclosure; SEQ ID NO 9; 113pp; English.

XX The invention relates to a chimeric protein comprising zinc finger

CC domains and a heterologous protein transduction domain for translocating

CC the protein across a cellular membrane. The invention also relates to a

CC pharmaceutical composition comprising the chimeric protein and a carrier,

CC a nucleic acid comprising a coding sequence that encodes a polypeptide

CC that comprises a zinc finger domain and a heterologous protein

CC transduction domain, located N-terminal to the zinc finger domain, a

CC method of altering gene expression in a cell of a subject, a method of

CC altering gene expression in a eukaryotic culture cell, a eukaryotic cell

CC that contains an exogenous polypeptide, but not a nucleic acid that

CC encodes the exogenous polypeptide, where the exogenous polypeptide

CC comprises zinc finger domains and a protein transduction domain that is

CC heterologous to the DNA binding domain and a method for preparing a

CC transducible DNA binding polypeptide. The chimeric protein is useful in

CC preparing a composition for treating a subject having or being suspected

CC of having a neoplastic, inflammatory or angiogenesis-based disorder. This

CC sequence represents the yeast UME6 repression domain used in the scope of

CC the invention.

XX

SQ Sequence 90 AA;

Query Match 52.6%; Score 40; DB 9; Length 90;

Best Local Similarity 64.3%; Pred. No. 31;

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHNSSE 13

|||||

64 VLRPILIRHNSSE 77

DE

RESULT 51

AAG91827

ID AAG91827 standard; protein; 115 AA.

XX

AC AAG91827;

XX

DT 26-SEP-2001 (first entry)

XX

DE C glutamicum protein fragment SEQ ID NO: 5581.

XX

KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX

OS Corynebacterium glutamicum.

XX

PN EP1108790-A2.

XX

PD 20-JUN-2001.

XX

PF 18-DEC-2000; 2000EP-00127688.

XX

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX

PA (KYOWA) KYOWA HAKKO KOGYO KK.

XX

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX

DR WPI; 2001-376931/40.

DR N-PSDB; AAH67046.

XX

PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

XX

PS Claim 17; SEQ ID NO 5581; 246pp + Sequence Listing; English.

XX

CC The present invention provides a number of nucleotide and protein

CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

are useful for identifying the mutation point of a gene derived from a

mutant of corynebacterium bacterium, measuring expression amount and analysing

the expression profile or expression pattern of a gene derived from

Corynebacterium bacterium, and identifying a homologue of a gene derived from

corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

acids, nucleic acids, vitamins, saccharides and organic acids,

particularly L-lysine. The present sequence is a protein described in the

exemplification of the invention. Note: The sequence data for this patent

did not form part of the printed specification, but was obtained in

electronic format directly from the European Patent Office

XX

SQ Sequence 115 AA;

Query Match 52.6%; Score 40; DB 4; Length 115;

Best Local Similarity 72.7%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 PSRVHNSSE 15

|||||

43 PLRVHNSSE 53

DE

RESULT 52

ADL33681

ID ADL33681 standard; protein; 210 AA.

XX

AC ADL33681;

XX

DT 20-MAY-2004 (first entry)

XX

DE Zinc finger domain-containing protein #1.

XX

KW modified cell; artificial transcription factor; stress resistance;

KW phenotypic trait alteration; zinc finger domain.

XX

OS Unidentified.

XX

PN WO2003048345-A1.

XX

PD 12-JUN-2003.

XX

PF 07-DEC-2002; 2002WO-KR002309.

XX

PR 07-DEC-2001; 2001US-0338441P.

PR 26-APR-2002; 2002US-0376053P.

PR 02-AUG-2002; 2002US-0400904P.

PR 05-AUG-2002; 2002US-0401089P.

XX

PA (TOOL-) TOOLGEN INC.

XX

PI Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;

PI Jang Y;

XX

DR WPI; 2003-513760/48.

XX

PT New modified cell comprising a heterologous nucleic acid encoding an

PT artificial transcription factor that confers stress resistance, useful

PT for altering a phenotypic trait of a cell or organism.

XX

PS Example 3; Page 99; 169pp; English.

XX

CC The invention comprises a modified cell containing a heterologous nucleic

CC acid encoding an artificial transcription factor that confers stress

CC resistance to the modified cell. The modified cell of the invention is

CC useful for altering a phenotypic trait of a cell or organism. The present

CC amino acid sequence represents a protein which contains zinc finger

CC domains.

XX

SQ Sequence 210 AA;

Query Match 52.6%; Score 40; DB 7; Length 210;

Best Local Similarity 64.3%; Pred. No. 83;

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

ADL33682
ID ADL33682 standard; protein; 238 AA.
XX
AC ADL33682;
XX
DT 20-MAY-2004 (first entry)
XX
DE Zinc finger domain-containing protein #2.
XX
KW modified cell; artificial transcription factor; stress resistance;
KW phenotypic trait alteration; zinc finger domain.
XX
OS Unidentified.
XX
PN WO2003048345-A1.
XX
PD 12-JUN-2003.
XX
PF 07-DEC-2002; 2002WO-KR002309.
XX
PR 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
XX
PA (TOOL-) TOOLGEN INC.
XX
PI Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
XX
DR WPI; 2003-513760/48.
XX
XX New modified cell comprising a heterologous nucleic acid encoding an
XX artificial transcription factor that confers stress resistance, useful
XX for altering a phenotypic trait of a cell or organism.
XX
PS Example 3; Page 99; 169pp; English.
XX
CC The invention comprises a modified cell containing a heterologous nucleic
XX acid encoding an artificial transcription factor that confers stress
XX resistance to the modified cell. The modified cell of the invention is
XX useful for altering a phenotypic trait of a cell or organism. The present
XX amino acid sequence represents a protein which contains zinc finger
XX domains.
XX
SQ Sequence 238 AA;
Query Match 52.6%; Score 40; DB 7; Length 238;
Best Local Similarity 64.3%; Pred. NO. 96;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Qy 2 VLRP--SRVHNSSE 13
Db 212 VLRPILLRHNSSE 225
RESULT 55
ADL33684
ID ADL33684 standard; protein; 238 AA.
XX
AC ADL33684;
XX
DT 20-MAY-2004 (first entry)
XX
DE Zinc finger domain-containing protein #4.
XX
KW modified cell; artificial transcription factor; stress resistance;
KW phenotypic trait alteration; zinc finger domain.
XX
OS Unidentified.
XX
PN WO2003048345-A1.
XX

Qy 2 VLRP--SRVHNSSE 13
Db 184 VLRPILLRHNSSE 197
RESULT 53
ADM72637
ID ADM72637 standard; protein; 210 AA.
XX
AC ADM72637;
XX
DT 03-JUN-2004 (first entry)
XX
DE Amino acid sequence of ZFD of K1.
XX
KW ZFP; zinc finger protein; gene evaluation; transcription factor;
KW zinc finger domain; zinc finger domain; ZFD.
XX
OS Synthetic.
XX
PN WO2004022575-A2.
XX
PD 18-MAR-2004.
XX
PF 05-SEP-2003; 2003WO-KR001827.
XX
PR 05-SEP-2002; 2002US-0408862P.
PR 07-MAR-2003; 2003US-0453111P.
XX
PA (TOOL-) TOOLGEN INC.
XX
PI Kim J, Lee D, Park J, Kim Y;
XX
DR WPI; 2004-248440/23.
XX
XX Evaluating cellular genes by expressing nucleic acid encoding different
XX artificial, chimeric transcription factor in cells and identifying set of
XX genes altered by transcription factors.
XX
PS Example 2; SEQ ID NO 9; 124pp; English.
XX
CC The invention relates to evaluating several cellular genes. The method
XX involves providing several cells containing nucleic acid encoding
XX different artificial, chimeric transcription factor, expressing nucleic
XX acids in cells, evaluating expression of several genes in cells to
XX provide expression information and identifying from expression
XX information, a set of two or more genes whose expression is altered by
XX first and second transcription factors. The nucleic acids encoding the
XX different artificial, chimeric transcription factors comprise a
XX preselected set of nucleic acids, where each member of the preselected
XX set of nucleic acids encodes an artificial, chimeric transcription factor
XX that can alter a cellular phenotype. Each artificial, chimeric
XX transcription factor comprises a first zinc finger domain, and further
XX comprises a second and third zinc finger domains. The first zinc finger
XX domain is a naturally occurring zinc finger domain. The first zinc finger
XX domain is a mammalian zinc finger domain. Each cell of several cells is
XX an animal or human cell. The method is useful for evaluating several
XX cellular genes. Sequences ADM72637-ADM72647 represents the amino acid
XX sequences of zinc finger domains (ZFD) of K1 to K11.
XX
SQ Sequence 210 AA;
Query Match 52.6%; Score 40; DB 8; Length 210;
Best Local Similarity 64.3%; Pred. NO. 83;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Qy 2 VLRP--SRVHNSSE 13
Db 184 VLRPILLRHNSSE 197
RESULT 54


```

PD 12-JUN-2003.
XX
PF 07-DEC-2002; 2002WO-KR002309.
XX
XX 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
XX (TOOL-) TOOLGEN INC.
XX Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
XX WPI; 2003-513760/48.
DR WPI; 2003-513760/48.
XX
XX New modified cell comprising a heterologous nucleic acid encoding an
PT artificial transcription factor that confers stress resistance, useful
PT for altering a phenotypic trait of a cell or organism.
XX
PS Example 3; SEQ ID NO 224; 169pp; English.
XX
XX The invention comprises a modified cell containing a heterologous nucleic
CC acid encoding an artificial transcription factor that confers stress
CC resistance to the modified cell. The modified cell of the invention is
CC useful for altering a phenotypic trait of a cell or organism. The present
CC amino acid sequence represents a protein which contains zinc finger
CC domains.
XX
XX Sequence 238 AA;
DR WPI; 2003-513760/48.
XX
XX Query Match 52.6%; Score 40; DB 7; Length 238;
XX Best Local Similarity 64.3%; Pred. No. 96;
XX Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
XX
QY 2 VLRP--SRVHNSSE 13
DB ||||| :||||:
212 VLRPILLRIHNSQ 225

RESULT 56
ADL33683
ID ADL33683 standard; protein; 238 AA.
XX
AC ADL33683;
XX
XX 20-MAY-2004 (first entry)
XX
XX Zinc finger domain-containing protein #3.
XX
XX modified cell; artificial transcription factor; stress resistance;
XX phenotypic trait alteration; zinc finger domain.
XX
XX Unidentified.
XX
XX WO2003048345-A1.
XX
XX 12-JUN-2003.
XX
XX 07-DEC-2002; 2002WO-KR002309.
XX
XX 07-DEC-2001; 2001US-0338441P.
PR 26-APR-2002; 2002US-0376053P.
PR 02-AUG-2002; 2002US-0400904P.
PR 05-AUG-2002; 2002US-0401089P.
XX (TOOL-) TOOLGEN INC.
XX Kim J, Park K, Lee D, Seol W, Lee H, Lee S, Yang H, Lee Y;
PI Jang Y;
XX WPI; 2003-513760/48.
XX

```

```

PT New modified cell comprising a heterologous nucleic acid encoding an
PT artificial transcription factor that confers stress resistance, useful
PT for altering a phenotypic trait of a cell or organism.
XX
XX Example 3; SEQ ID NO 223; 169pp; English.
XX
XX The invention comprises a modified cell containing a heterologous nucleic
CC acid encoding an artificial transcription factor that confers stress
CC resistance to the modified cell. The modified cell of the invention is
CC useful for altering a phenotypic trait of a cell or organism. The present
CC amino acid sequence represents a protein which contains zinc finger
CC domains.
XX
XX Sequence 238 AA;
DR WPI; 2003-513760/48.
XX
XX Query Match 52.6%; Score 40; DB 7; Length 238;
XX Best Local Similarity 64.3%; Pred. No. 96;
XX Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
XX
QY 2 VLRP--SRVHNSSE 13
DB ||||| :||||:
212 VLRPILLRIHNSQ 225

RESULT 57
ADM72638
ID ADM72638 standard; protein; 238 AA.
XX
AC ADM72638;
XX
XX 03-JUN-2004 (first entry)
XX
XX Amino acid sequence of ZFD of K2.
XX
XX ZFP; zinc finger protein; gene evaluation; transcription factor;
XX zinc finger domain; zinc finger domain; ZFD.
XX
XX Synthetic.
XX
XX WO2004022575-A2.
XX
XX 18-MAR-2004.
XX
XX 05-SEP-2003; 2003WO-KR001827.
XX
XX 05-SEP-2002; 2002US-0408862P.
PR 07-MAR-2003; 2003US-0453111P.
XX (TOOL-) TOOLGEN INC.
XX Kim J, Lee D, Park J, Kim Y;
XX WPI; 2004-248440/23.
XX
XX Evaluating cellular genes by expressing nucleic acid encoding different
PT artificial, chimeric transcription factor in cells and identifying set of
PT genes altered by transcription factors.
XX
XX Example 2; SEQ ID NO 10; 124pp; English.
XX
XX The invention relates to evaluating several cellular genes. The method
CC involves providing several cells containing nucleic acid encoding
CC different artificial, chimeric transcription factor, expressing nucleic
CC acids in cells, evaluating expression of several genes in cells to
CC provide expression information and identifying from expression
CC information, a set of two or more genes whose expression is altered by
CC first and second transcription factors. The nucleic acids encoding the
CC different artificial, chimeric transcription factors comprise a
CC preselected set of nucleic acids, where each member of the preselected
CC set of nucleic acids encodes an artificial, chimeric transcription factor
CC that can alter a cellular phenotype. Each artificial, chimeric
CC transcription factor comprises a first zinc finger domain, and further
CC comprises a second and third zinc finger domains. The first zinc finger

```

CC domain is a naturally occurring zinc finger domain. The first zinc finger
CC domain is a mammalian zinc finger domain. Each cell of several cells is
CC an animal or human cell. The method is useful for evaluating several
CC cellular genes. Sequences ADM72637-ADM72647 represents the amino acid
CC sequences of zinc finger domains (ZFD) of K1 to K11.
XX
SQ Sequence 238 AA;

Query Match 52.6%; Score 40; DB 8; Length 238;
Best Local Similarity 64.3%; Pred. No. 96;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 2 VLRP--SRVHNSSE 13
|||||:|||||:
Db 212 VLRPILIRHNSQ 225

RESULT 58
ADM72639
ID ADM72639 standard; protein; 238 AA.

XX AC ADM72639;

XX DT 03-JUN-2004 (first entry)

XX DE Amino acid sequence of ZFD of K3.

XX KW ZFP; zinc finger protein; gene evaluation; transcription factor;
XX KW zinc finger domain; zinc finger domain; ZFD.

XX OS Synthetic.

XX PN WO2004022575-A2.

XX PD 18-MAR-2004.

XX PF 05-SEP-2003; 2003WO-KR001827.

XX PR 05-SEP-2002; 2002US-0408862P.

XX PR 07-MAR-2003; 2003US-0453111P.

XX PA (TOOL-) TOOLGEN INC.

XX PI Kim J, Lee D, Park J, Kim Y;

XX PS WPI; 2004-248440/23.

XX DR Evaluating cellular genes by expressing nucleic acid encoding different
XX PT artificial, chimeric transcription factor in cells and identifying set of
XX PT genes altered by transcription factors.

XX PS Example 2; SEQ ID NO 11; 124pp; English.

XX CC The invention relates to evaluating several cellular genes. The method
XX CC involves providing several cells containing nucleic acid encoding
XX CC different artificial, chimeric transcription factor, expressing nucleic
XX CC acids in cells, evaluating expression of several genes in cells to
XX CC provide expression information and identifying from expression
XX CC information, a set of two or more genes whose expression is altered by
XX CC first and second transcription factors. The nucleic acids encoding the
XX CC different artificial, chimeric transcription factors comprise a
XX CC preselected set of nucleic acids, where each member of the preselected
XX CC set of nucleic acids encodes an artificial, chimeric transcription factor
XX CC that can alter a cellular phenotype. Each artificial, chimeric
XX CC transcription factor comprises a first zinc finger domain, and further
XX CC comprises a second and third zinc finger domains. The first zinc finger
XX CC domain is a naturally occurring zinc finger domain. The first zinc finger
XX CC domain is a mammalian zinc finger domain. Each cell of several cells is
XX CC an animal or human cell. The method is useful for evaluating several
XX CC cellular genes. Sequences ADM72637-ADM72647 represents the amino acid
XX CC sequences of zinc finger domains (ZFD) of K1 to K11.

XX SQ Sequence 238 AA;

Query Match 52.6%; Score 40; DB 8; Length 238;
Best Local Similarity 64.3%; Pred. No. 96;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 2 VLRP--SRVHNSSE 13
|||||:|||||:
Db 212 VLRPILIRHNSQ 225

RESULT 59

ADM72640
ID ADM72640 standard; protein; 238 AA.

XX AC ADM72640;

XX DT 03-JUN-2004 (first entry)

XX DE Amino acid sequence of ZFD of K4.

XX KW ZFP; zinc finger protein; gene evaluation; transcription factor;
XX KW zinc finger domain; zinc finger domain; ZFD.

XX OS Synthetic.

XX PN WO2004022575-A2.

XX PD 18-MAR-2004.

XX PF 05-SEP-2003; 2003WO-KR001827.

XX PR 05-SEP-2002; 2002US-0408862P.

XX PR 07-MAR-2003; 2003US-0453111P.

XX PA (TOOL-) TOOLGEN INC.

XX PI Kim J, Lee D, Park J, Kim Y;

XX PS WPI; 2004-248440/23.

XX PT Evaluating cellular genes by expressing nucleic acid encoding different
XX PT artificial, chimeric transcription factor in cells and identifying set of
XX PT genes altered by transcription factors.

XX PS Example 2; SEQ ID NO 12; 124pp; English.

XX CC The invention relates to evaluating several cellular genes. The method
XX CC involves providing several cells containing nucleic acid encoding
XX CC different artificial, chimeric transcription factor, expressing nucleic
XX CC acids in cells, evaluating expression of several genes in cells to
XX CC provide expression information and identifying from expression
XX CC information, a set of two or more genes whose expression is altered by
XX CC first and second transcription factors. The nucleic acids encoding the
XX CC different artificial, chimeric transcription factors comprise a
XX CC preselected set of nucleic acids, where each member of the preselected
XX CC set of nucleic acids encodes an artificial, chimeric transcription factor
XX CC that can alter a cellular phenotype. Each artificial, chimeric
XX CC transcription factor comprises a first zinc finger domain, and further
XX CC comprises a second and third zinc finger domains. The first zinc finger
XX CC domain is a naturally occurring zinc finger domain. The first zinc finger
XX CC domain is a mammalian zinc finger domain. Each cell of several cells is
XX CC an animal or human cell. The method is useful for evaluating several
XX CC cellular genes. Sequences ADM72637-ADM72647 represents the amino acid
XX CC sequences of zinc finger domains (ZFD) of K1 to K11.

XX SQ Sequence 238 AA;

Query Match 52.6%; Score 40; DB 8; Length 238;
Best Local Similarity 64.3%; Pred. No. 96;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 2 VLRP--SRVHNSSE 13
|||||:|||||:

Mon Mar 13 10:54:34 2006

Search completed: March 11, 2006, 12:18:16
Job time : 161.132 secs

Db 212 VLRPILRHSEQ 225

RESULT 60
ADR10155
ID ADR10155 standard; protein; 310 AA.
XX AC ADR10155;
XX DT 04-NOV-2004 (first entry)
XX DE Human protein useful for treating neurological disease Seq 3661.
XX KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquilliser.
XX OS Homo sapiens.
XX PN EPI447413-A2.
XX PD 18-AUG-2004.
XX PF 12-FEB-2004; 2004EP-00003145.
XX PR 14-FEB-2003; 2003JP-00102207.
XX PR 09-MAY-2003; 2003JP-00131452.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX DR WPI; 2004-593265/57.
XX DR N-PSDB; ADR08199.
XX PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
XX PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX PS Claim 1; SEQ ID NO 3661; 2686pp; English.
XX CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquilliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
XX office.
SQ Sequence 310 AA;

Query Match 52.6%; Score 40; DB 8; Length 310;
Best Local Similarity 46.7%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSENT 15
:::|:|:
DB 23 LILRFQMQHPTTENT 37

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:19:28 ; Search time 23.6842 Seconds

(without alignments)
60.937 Million cell updates/sec

Title: US-10-774-242A-1

Perfect score: 76

Sequence: 1 IVLRPSRVHSEENT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

PIR_80:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	759	2 I38593	fibroblast activat
2	42	55.3	444	2 AD2031	hypothetical prote
3	41	53.9	740	2 T26140	hypothetical prote
4	41	53.9	937	2 T04194	hypothetical prote
5	40	52.6	532	2 T14335	protein kinase, ca
6	40	52.6	836	2 S61570	regulatory protein
7	39	51.3	333	2 D85068	D123-like protein
8	39	51.3	917	1 RDMNH	nitrate reductase
9	38.5	50.7	272	2 B69387	fumarase (fum-2) h
10	38	50.0	62	2 D71571	hypothetical prote
11	38	50.0	259	2 S74287	hypothetical prote
12	38	50.0	368	1 S64487	GTP-binding protei
13	38	50.0	928	2 T04192	hypothetical prote
14	38	50.0	948	2 T37816	probable peroxin-6
15	38	50.0	1278	2 A71609	probable secreted
16	37	48.7	39	2 D40984	finger protein zfa
17	37	48.7	128	2 S77673	streptokinase A (S
18	37	48.7	202	2 T01510	hypothetical prote
19	37	48.7	272	2 S61117	probable membrane
20	37	48.7	301	2 T23403	hypothetical prote
21	37	48.7	352	2 B43649	avirulence protein
22	37	48.7	375	2 C81960	hypothetical prote
23	37	48.7	388	2 T32994	hypothetical prote
24	37	48.7	418	2 A46226	somatostatin recep
25	37	48.7	491	2 H84379	4-hydroxybutyrate
26	37	48.7	492	2 G65076	hypothetical prote
27	37	48.7	492	2 F91102	probable coenzyme
28	37	48.7	492	2 B85948	probable coenzyme
29	37	48.7	544	2 A47726	dis1-suppressing p

30	37	48.7	584	2 AC3547	ribose transport A
31	37	48.7	732	2 B84902	hypothetical prote
32	37	48.7	737	2 AC0403	catalase (EC 1.11.
33	37	48.7	758	2 T31994	hypothetical prote
34	37	48.7	808	1 T04982	dynamitin-like prote
35	37	48.7	821	2 D64119	glycogen phosphory
36	37	48.7	847	2 T14470	receptor-like kina
37	36	47.4	196	2 S74756	hypothetical prote
38	36	47.4	220	2 T49580	probable oligomyci
39	36	47.4	269	2 D82060	conserved hypother
40	36	47.4	282	2 D97280	fumarate hydratase
41	36	47.4	295	2 D95107	hypothetical prote
42	36	47.4	295	2 F97975	hypothetical prote
43	36	47.4	305	2 E97160	probable pseudouri
44	36	47.4	311	1 DTECC	aspartate carbamoy
45	36	47.4	311	2 G86122	hypothetical prote
46	36	47.4	311	2 F91281	aspartate carbamoy
47	36	47.4	392	2 T34873	hypothetical prote
48	36	47.4	429	1 S44778	transcription fact
49	36	47.4	467	1 A30042	transcription fact
50	36	47.4	552	2 H97289	dihydroxy-acid deh
51	36	47.4	578	2 G75622	probable oligopept
52	36	47.4	595	2 T31163	hypothetical prote
53	36	47.4	620	2 S84304	hypothetical prote
54	36	47.4	634	2 T27881	hypothetical prote
55	36	47.4	691	2 T34573	probable membrane
56	36	47.4	733	2 S56277	hypothetical prote
57	36	47.4	989	2 T45283	growth polarity ma
58	36	47.4	1066	2 T45283	staurosporine targ
59	36	47.4	1066	2 T41099	SSD1 protein - yea
60	36	47.4	1250	2 A39578	F7H2.12 protein -
61	36	47.4	1366	2 B86292	hypothetical prote
62	36	47.4	1384	2 T26656	probable membrane
63	36	47.4	1679	2 S49802	hypothetical prote
64	35.5	46.7	310	2 AC1859	hypothetical prote
65	35	46.1	118	2 E84062	hypothetical prote
66	35	46.1	150	2 G87424	hypothetical prote
67	35	46.1	173	1 NKVCCU	core protein p20 -
68	35	46.1	180	2 A70439	probable dCTP deam
69	35	46.1	230	2 C90266	heterodisulfide re
70	35	46.1	232	2 C99938	conserved hypother
71	35	46.1	271	2 B83682	urase accessory p
72	35	46.1	297	2 F70201	conserved hypother
73	35	46.1	298	2 E69109	conserved hypother
74	35	46.1	312	2 C70901	probable fnt prote
75	35	46.1	314	2 S55210	meiotic recombinat
76	35	46.1	333	2 F84166	LPS glycoasyltransf
77	35	46.1	335	2 S86939	hypothetical prote
78	35	46.1	387	2 AF2115	hypothetical prote
79	35	46.1	405	2 T23370	hypothetical prote
80	35	46.1	428	2 A44021	somatostatin recep
81	35	46.1	428	2 S30508	probable G protein
82	35	46.1	521	2 S06218	colicin E1 - Shige
83	35	46.1	522	1 IRECI	colicin E1 - Esche
84	35	46.1	522	2 I59270	colicin E1 - Esche
85	35	46.1	588	2 E89945	aspartyl-tRNA synt
86	35	46.1	595	2 D83806	aspartyl-tRNA synt
87	35	46.1	601	2 E87028	pyruvate, phosphat
88	35	46.1	663	1 A54075	arachidonate 12-li
89	35	46.1	707	2 T09340	hypothetical prote
90	35	46.1	765	2 C42632	cell adhesion mole
91	35	46.1	812	2 B42632	cell adhesion mole
92	35	46.1	854	2 S33558	unc-33 protein - C
93	35	46.1	865	2 T40170	hypothetical prote
94	35	46.1	932	2 A42632	cell adhesion mole
95	35	46.1	986	2 B81675	polymorphic membra
96	35	46.1	1189	2 T42726	guanine nucleotide
97	35	46.1	1279	2 E64709	type IIS restricti
98	35	46.1	1703	2 S15047	SNF2 protein - yea
99	35	46.1	1969	2 T08875	histidine kinase h
100	35	46.1	2573	2 D71614	hypothetical prote
101	34.5	45.4	835	2 I55603	reduced hepatic gl
102	34	44.7	158	2 S49918	hypothetical prote

103	34	44.7	158	2	AD3583	riboflavin synthas	176	34	44.7	2150	1	S27802	zinc finger protei
104	34	44.7	170	2	T17286	hypothetical prote	177	34	44.7	2150	2	T19450	hypothetical prote
105	34	44.7	174	2	T17650	hypothetical prote	178	34	44.7	2161	1	A45389	genome polyprotein
106	34	44.7	181	1	C69847	ribosomal-protein-	179	34	44.7	2599	2	A96616	unknown protein F1
107	34	44.7	194	2	I51310	beta 8 integrin -	180	33.5	44.1	161	2	AC1984	hypothetical prote
108	34	44.7	200	2	A82933	hypothetical prote	181	33.5	44.1	408	2	SS1467	hypothetical prote
109	34	44.7	212	2	C64583	conserved hypotet	182	33.5	44.1	724	1	JQ1622	glycoprotein H pre
110	34	44.7	212	2	F71928	hypothetical prote	183	33	43.4	89	2	S38676	MHC class II histo
111	34	44.7	218	2	D72763	hypothetical prote	184	33	43.4	89	2	AB1310	ribosomal protein
112	34	44.7	244	2	T51481	hypothetical prote	185	33	43.4	104	2	S56789	hypothetical prote
113	34	44.7	265	2	T29973	hypothetical prote	186	33	43.4	113	2	AG1918	hypothetical prote
114	34	44.7	273	2	B69883	conserved hypotet	187	33	43.4	120	2	AH1329	cell-division prot
115	34	44.7	301	2	F86749	pseudouridine synt	188	33	43.4	120	2	AH1700	cell-division prot
116	34	44.7	311	1	S48560	purine-nucleoside	189	33	43.4	128	2	S77677	streptokinase A (E
117	34	44.7	327	2	B69856	oligopeptide trans	190	33	43.4	128	2	S77683	streptokinase A (E
118	34	44.7	327	2	D71278	hypothetical prote	191	33	43.4	128	2	S77687	streptokinase A (E
119	34	44.7	366	2	S76662	S-adenosylmethioni	192	33	43.4	128	2	S77682	streptokinase A (E
120	34	44.7	377	2	A47380	RING finger-contai	193	33	43.4	128	2	S77686	streptokinase A (E
121	34	44.7	400	2	F65013	nucleoside transpo	194	33	43.4	128	2	S77684	streptokinase A (E
122	34	44.7	402	1	F48552	protein kinase (EC	195	33	43.4	128	2	S77681	streptokinase A (E
123	34	44.7	403	2	S50944	probable membrane	196	33	43.4	128	2	S77678	streptokinase A (E
124	34	44.7	414	2	C98041	conserved hypotet	197	33	43.4	128	2	S77676	streptokinase A (E
125	34	44.7	440	2	I50213	protein-tyrosine-p	198	33	43.4	128	2	S77685	streptokinase A (E
126	34	44.7	458	2	S24457	hypothetical prote	199	33	43.4	128	2	S77675	streptokinase A (E
127	34	44.7	468	2	T50992	origin recognition	200	33	43.4	128	2	S77675	streptokinase A (E
128	34	44.7	497	1	S43532	protein-tyrosine k	201	33	43.4	136	2	T45360	photosystem II pro
129	34	44.7	498	2	T45465	acetate CoA-transf	202	33	43.4	137	2	S53025	hypothetical prote
130	34	44.7	498	2	F96622	hypothetical prote	203	33	43.4	137	2	S53025	hypothetical prote
131	34	44.7	557	2	T15250	hypothetical prote	204	33	43.4	137	2	S53025	hypothetical prote
132	34	44.7	575	2	T28230	ORF MSV069 probabl	205	33	43.4	137	2	S53025	hypothetical prote
133	34	44.7	577	2	A64597	hypothetical prote	206	33	43.4	137	2	S53025	hypothetical prote
134	34	44.7	577	2	F86281	hypothetical prote	207	33	43.4	137	2	S53025	hypothetical prote
135	34	44.7	579	2	C71916	aspartate-tRNA lig	208	33	43.4	137	2	S53025	hypothetical prote
136	34	44.7	581	2	JC7086	F2D10 protein - hu	209	33	43.4	137	2	S53025	hypothetical prote
137	34	44.7	583	2	G81412	aspartate-tRNA lig	210	33	43.4	137	2	S53025	hypothetical prote
138	34	44.7	592	2	D69591	aspartate-tRNA lig	211	33	43.4	137	2	S53025	hypothetical prote
139	34	44.7	594	2	T43246	amino acid transpo	212	33	43.4	137	2	S53025	hypothetical prote
140	34	44.7	597	2	B75556	hypothetical prote	213	33	43.4	137	2	S53025	hypothetical prote
141	34	44.7	613	2	S66977	hypothetical prote	214	33	43.4	137	2	S53025	hypothetical prote
142	34	44.7	654	2	B84690	hypothetical prote	215	33	43.4	137	2	S53025	hypothetical prote
143	34	44.7	671	2	S51599	Om(2D) protein - f	216	33	43.4	137	2	S53025	hypothetical prote
144	34	44.7	687	2	T26332	hypothetical prote	217	33	43.4	137	2	S53025	hypothetical prote
145	34	44.7	689	2	T26331	hypothetical prote	218	33	43.4	137	2	S53025	hypothetical prote
146	34	44.7	691	2	F86132	hypothetical prote	219	33	43.4	137	2	S53025	hypothetical prote
147	34	44.7	708	2	T19474	hypothetical prote	220	33	43.4	137	2	S53025	hypothetical prote
148	34	44.7	728	2	JC5134	Glycosyltransferas	221	33	43.4	137	2	S53025	hypothetical prote
149	34	44.7	732	2	S73089	glycosyltransferas	222	33	43.4	137	2	S53025	hypothetical prote
150	34	44.7	733	2	C99940	hypothetical prote	223	33	43.4	137	2	S53025	hypothetical prote
151	34	44.7	772	2	T26330	hypothetical prote	224	33	43.4	137	2	S53025	hypothetical prote
152	34	44.7	782	1	FAHUP	galactin precursor	225	33	43.4	137	2	S53025	hypothetical prote
153	34	44.7	796	2	A38992	cadherin-11 precu	226	33	43.4	137	2	S53025	hypothetical prote
154	34	44.7	796	2	I48277	cadherin-11 - mous	227	33	43.4	137	2	S53025	hypothetical prote
155	34	44.7	796	2	I49556	cadherin-11 - mous	228	33	43.4	137	2	S53025	hypothetical prote
156	34	44.7	796	2	A53584	OB-cadherin precu	229	33	43.4	137	2	S53025	hypothetical prote
157	34	44.7	809	2	E71660	hypothetical prote	230	33	43.4	137	2	S53025	hypothetical prote
158	34	44.7	829	2	T46556	DNA topoisomerase	231	33	43.4	137	2	S53025	hypothetical prote
159	34	44.7	830	2	G36900	DNA gyrase (topois	232	33	43.4	137	2	S53025	hypothetical prote
160	34	44.7	833	1	A31593	heat shock transcr	233	33	43.4	137	2	S53025	hypothetical prote
161	34	44.7	899	2	C71608	origin recognition	234	33	43.4	137	2	S53025	hypothetical prote
162	34	44.7	904	2	I38757	homolog of Drosoph	235	33	43.4	137	2	S53025	hypothetical prote
163	34	44.7	916	2	H72372	exonuclease ABC c	236	33	43.4	137	2	S53025	hypothetical prote
164	34	44.7	919	1	PNZ21P	H+-exporting ATPas	237	33	43.4	137	2	S53025	hypothetical prote
165	34	44.7	926	2	I38756	homolog of Drosoph	238	33	43.4	137	2	S53025	hypothetical prote
166	34	44.7	1159	2	T02866	hypothetical prote	239	33	43.4	137	2	S53025	hypothetical prote
167	34	44.7	1186	2	S72229	meiotic recombination	240	33	43.4	137	2	S53025	hypothetical prote
168	34	44.7	1201	2	H86434	protein F1F8.21 (241	33	43.4	137	2	S53025	hypothetical prote
169	34	44.7	1252	2	D71810	probable type II D	242	33	43.4	137	2	S53025	hypothetical prote
170	34	44.7	1389	2	T47796	ABC transporter-li	243	33	43.4	137	2	S53025	hypothetical prote
171	34	44.7	1465	2	A70199	hypothetical prote	244	33	43.4	137	2	S53025	hypothetical prote
172	34	44.7	1472	2	B54774	ATP binding casset	245	33	43.4	137	2	S53025	hypothetical prote
173	34	44.7	1529	2	A59189	ATP-binding casset	246	33	43.4	137	2	S53025	hypothetical prote
174	34	44.7	1541	2	AG2474	heterocyst glycoli	247	33	43.4	137	2	S53025	hypothetical prote
175	34	44.7	1872	2	T00339	hypothetical prote	248	33	43.4	137	2	S53025	hypothetical prote

249 33 43.4 462 2 AC2605 transcription regu
250 33 43.4 465 2 S01548 protochlorophyllid
251 33 43.4 470 2 B37407 probable coenzyme
252 33 43.4 491 2 AH3573 glucose-6-phosphat
253 33 43.4 496 2 AB3368 acetate CoA-transf
254 33 43.4 497 2 B82966 probable coenzyme
255 33 43.4 498 2 AC2625 coenzyme A transfe
256 33 43.4 514 2 B87711 coenzyme A transfe
257 33 43.4 529 2 T08684 hypohetical prote
258 33 43.4 577 2 A57469 CMP-N-acetylneuram
259 33 43.4 585 2 B75265 probable acid-CoA
260 33 43.4 596 2 A96710 hypohetical prote
261 33 43.4 604 2 T37870 RNA-binding / Ran
262 33 43.4 619 2 C97114 unknown protein T6
263 33 43.4 627 2 T27123 hypohetical prote
264 33 43.4 629 2 E64046 mismatch repair pr
265 33 43.4 633 2 J01242 viral replicase 2
266 33 43.4 639 2 T50793 hypohetical prote
267 33 43.4 663 1 A35087 arachidonate 12-li
268 33 43.4 663 1 S30051 arachidonate 12-li
269 33 43.4 663 2 I52462 arachidonate 12-li
270 33 43.4 712 2 S23650 retrovirus-related
271 33 43.4 736 2 T00313 catalase (EC 1.11.
272 33 43.4 759 2 S05044 hypohetical prote
273 33 43.4 763 2 T27937 hypohetical prote
274 33 43.4 764 1 BBH0 integrin beta-8 ch
275 33 43.4 768 2 B41029 integrin beta-8 ch
276 33 43.4 769 2 A1029 integrin beta-8 ch
277 33 43.4 772 2 T08226 hypohetical prote
278 33 43.4 857 2 T20318 hypohetical prote
279 33 43.4 875 2 C81209 tsPA protein NMB03
280 33 43.4 878 2 A97854 aconitate hydratase
281 33 43.4 878 2 A71641 aconitate hydratase
282 33 43.4 921 2 D82513 malt regulatory pr
283 33 43.4 942 2 T19553 hypohetical prote
284 33 43.4 950 2 T25088 hypohetical prote
285 33 43.4 961 2 E86245 hypohetical prote
286 33 43.4 997 2 J06067 CCAAT-binding fact
287 33 43.4 1018 2 T40253 hypohetical prote
288 33 43.4 1026 2 A89696 protein f21H9.1 li
289 33 43.4 1072 2 A86827 hypohetical prote
290 33 43.4 1115 2 T09403 integrin alpha cha
291 33 43.4 1225 2 C84530 hypohetical prote
292 33 43.4 1247 2 T13131 nitric-oxide synth
293 33 43.4 1307 2 T30887 146D nuclear prote
294 33 43.4 1310 2 A81749 glycosidase homolo
295 33 43.4 1417 2 T21012 hypohetical prote
296 33 43.4 1466 2 T17138 CLIAA protein - ra
297 33 43.4 1467 2 T18411 latrophilin-1, bra
298 33 43.4 1471 2 T17149 CLIBA protein - ra
299 33 43.4 1471 2 F86218 protein f22O13.8
300 33 43.4 1472 2 T18413 latrophilin-1, bra
301 33 43.4 1496 2 B81416 glutamate synthase
302 33 43.4 1496 2 T40016 conserved hypohet
303 33 43.4 1510 2 T17145 CLIAA protein - ra
304 33 43.4 1515 2 T17156 CLIBB protein - ra
305 33 43.4 1520 2 G96634 glutamate synthase
306 33 43.4 1530 2 A81291 glutamate synthase
307 33 43.4 1536 1 R8BYS3 glutamate synthase
308 33 43.4 1606 2 T34073 regulatory protein
309 33 43.4 1616 2 T00713 paraneurin - chicke
310 33 43.4 1657 2 T25421 helicase homolog F
311 33 43.4 1696 2 T23617 hypohetical prote
312 33 43.4 1825 2 T52521 related to SEN1 pr
313 33 43.4 1840 2 T29091 transitin - chicke
314 33 43.4 1944 2 A59438 KIAA1424 protein
315 33 43.4 2037 2 S37178 fatty-acyl-CoA syn
316 33 43.4 2207 2 T24629 glutamate synthase
317 33 43.4 2493 2 S45734 probable membrane
318 33 43.4 2493 2 S45734 probable membrane
319 33 43.4 2493 2 S45734 probable membrane
320 32.5 42.8 325 2 A69184 UDP-glucose 4-epim
321 32.5 42.8 341 2 T33990 hypohetical prote

322 32.5 42.8 437 2 T42653 hypohetical prote
323 32.5 42.8 586 2 JC7277 Sarcophaga-derived
324 32.5 42.8 1695 2 T19823 hypohetical prote
325 32 42.1 78 2 T13257 hypohetical prote
326 32 42.1 79 2 H82705 hypohetical prote
327 32 42.1 88 2 T05438 hypohetical prote
328 32 42.1 106 2 A72598 probable lipoprote
329 32 42.1 111 2 AD0961 beta-galactosidase
330 32 42.1 113 4 A31199 hypohetical prote
331 32 42.1 123 2 T23448 hypohetical prote
332 32 42.1 129 2 F90293 hypohetical prote
333 32 42.1 142 2 B83638 hypohetical prote
334 32 42.1 153 2 S63255 hypohetical prote
335 32 42.1 165 2 A47148 reg I, regeneratin
336 32 42.1 169 2 H97772 hypohetical prote
337 32 42.1 202 2 B86447 nonstructural prot
338 32 42.1 204 1 MN283 hypohetical prote
339 32 42.1 216 2 F84292 probable finger pr
340 32 42.1 220 2 S45927 probable membrane
341 32 42.1 223 2 S48785 hypohetical prote
342 32 42.1 231 2 S55078 hypohetical prote
343 32 42.1 244 2 T28175 ABC transporter po
344 32 42.1 247 1 B71694 hypohetical prote
345 32 42.1 250 2 T00911 modulation protein
346 32 42.1 265 2 S34668 dihydroperoxide sy
347 32 42.1 267 2 C89818 probable transcript
348 32 42.1 271 2 A83335 glycosyl transfera
349 32 42.1 276 2 S70815 lacto-N-neotetraos
350 32 42.1 280 2 A81027 hypohetical prote
351 32 42.1 280 2 H71552 hypohetical prote
352 32 42.1 284 2 T04674 hypohetical prote
353 32 42.1 286 2 T38680 probable f0-f0 par
354 32 42.1 290 1 XYRBP arylamine N-acetyl
355 32 42.1 291 2 T45698 hypohetical prote
356 32 42.1 303 1 A42544 nonstructural prot
357 32 42.1 310 2 D71878 stress protein - H
358 32 42.1 315 2 E83369 probable transcript
359 32 42.1 325 2 D69620 electron transfer
360 32 42.1 326 2 AH2612 NADH-ubiquinone ox
361 32 42.1 328 2 A13541 NADH2 dehydrogenas
362 32 42.1 329 2 T45200 endoglucanase - Th
363 32 42.1 339 2 C72301 hypohetical prote
364 32 42.1 347 2 B72596 hypohetical prote
365 32 42.1 351 2 S12013 glucan endo-1,3-be
366 32 42.1 354 2 H69815 hypohetical prote
367 32 42.1 354 2 F97394 probable NADH-ubiq
368 32 42.1 358 2 C69348 hypohetical prote
369 32 42.1 359 2 T19009 hypohetical prote
370 32 42.1 360 2 T20853 hypohetical prote
371 32 42.1 361 2 AE2074 ferrichrome bindin
372 32 42.1 372 2 S69064 hypohetical prote
373 32 42.1 384 2 G56689 probable fructokin
374 32 42.1 386 2 T35438 probable secreted
375 32 42.1 386 2 S41497 thyrold hormone re
376 32 42.1 389 2 T43905 DNA gyrase chain B
377 32 42.1 390 2 AG3180 agrobacterium viru
378 32 42.1 390 2 F95350 VirB10 transembra
379 32 42.1 392 2 S72984 hypohetical prote
380 32 42.1 396 2 S68772 ggdef family prote
381 32 42.1 396 2 C97512 hypohetical prote
382 32 42.1 409 2 S60975 GGDEF family prote
383 32 42.1 415 2 AD2731 GGDEF family prote
384 32 42.1 417 1 S00845 hepsin (EC 3.4.21.
385 32 42.1 418 2 AB1170 ATP/GTP-binding pr
386 32 42.1 418 2 AD1527 ATP/GTP-binding pr
387 32 42.1 424 2 A50000 porphobilinogen sy
388 32 42.1 425 2 AF0055 probable exported
389 32 42.1 438 2 S70232 tYA protein - Yeas
390 32 42.1 439 2 A11388 late competence pr
391 32 42.1 442 2 JC5594 jerky gene protein
392 32 42.1 455 2 C95049 acetyl-CoA carboxy
393 32 42.1 455 2 A97920 biotin carboxylase
394 32 42.1 455 2 D96525 protein TIN15.19 (

C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AD2031
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shampo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 203-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <KUR>
A:Cross-references: UNIPROT:O8YW14; UNIPARC:UPI00000CE1F8; GB:BA0000019; PIDN:BA073501.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1802
C:Superfamily: Na+-ATP synthase chain J

Query Match 55.3%; Score 42; DB 2; Length 444;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEN 14
||:|:|
Db 430 RPTRIHYPEEN 440

RESULT 3
T26140
hypothetical protein W04A8.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26140
R:Kershaw, J.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20159
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-740 <WIL>
A:Cross-references: UNIPROT:Q9XUM0; UNIPARC:UPI000007842C; EMBL:Z82069; PIDN:CAB04906.1;
A:Experimental source: clone W04A8
C:Genetics:
A:Gene: CESP:W04A8.6
A:Map position: 1
A:Introns: 122/3; 194/3; 218/3; 266/1; 293/2; 378/2; 646/3; 697/2
C:Superfamily: Caenorhabditis elegans hypothetical protein W04A8.6

Query Match 53.9%; Score 41; DB 2; Length 740;
Best Local Similarity 46.2%; Pred. No. 29;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSEN 14
||:|:|
Db 204 MEKPSKKNKEN 216

RESULT 4
T04194
hypothetical protein T4F9.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04194
R:Bevan, M.; Wedler, H.; Wedler, R.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F. submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15260
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-937 <BEV>
A:Cross-references: UNIPROT:Q9T088; UNIPARC:UPI00000A40D1; EMBL:AL049523
A:Experimental source: cultivar Columbia; BAC clone T4F9
C:Genetics:

A:Map position: 4
A:Introns: 51/2; 122/2; 170/3; 190/3; 250/3; 375/3; 422/3; 479/3; 565/3; 617/2; 795/2; 8
A:Note: T4F9.50
C:Superfamily: ubiquitin C-terminal hydrolase Ubpl2

Query Match 53.9%; Score 41; DB 2; Length 937;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHNSSEN 14
||:|:|
Db 666 PSKIHNSSEN 675

RESULT 5
T14335
protein kinase, calcium-dependent (EC 2.7.1.1-) - carrot
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14335
R:Choi, J.H.
submitted to the EMBL Data Library, January 1997
A:Reference number: Z17988
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-532 <CHO>
A:Cross-references: UNIPROT:P28582; UNIPARC:UPI0000127447; EMBL:X56599
C:Genetics:
A:Note: DcPK431
C:Superfamily: calcium-dependent protein kinase; calmodulin repeat homology; protein kin
C:Keywords: calcium binding; EF hand; phosphotransferase
F:487-519/Domain: calmodulin repeat homology <EFH>

Query Match 52.6%; Score 40; DB 2; Length 532;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSEN 15
||:|:|
Db 56 LKPRQVHRPESNT 68

RESULT 6
S61570
regulatory protein NIM2 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: CAR80 protein; protein YD8142A.04c; protein YDR207c; transcription fa
C:Species: Saccharomyces cerevisiae
C:Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 05-Oct-2004
C:Accession: S61570; S45175; F53420; S51874
R:Murphy, L.; Harris, D.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61117
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <MUR>
A:Cross-references: UNIPROT:P39001; UNIPARC:UPI000005313A; EMBL:Z68194; NID:g1204148; PI
A:Experimental source: strain AB972
R:Kumeno, A.
submitted to the EMBL Data Library, November 1993
A:Reference number: S45175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100, 'G', 102-836 <KUM>
A:Cross-references: UNIPARC:UPI0000168D23; EMBL:D23663; NID:G432373; PIDN:BA004890.1; P
A:Experimental source: strain GRF88
R:Strich, R.; Surosky, R.T.; Steber, C.; Dubois, E.; Messenguy, F.; Esposito, R.E.
Genes Dev. 8, 796-810, 1994
A:Title: UME6 is a key regulator of nitrogen repression and meiotic development.
A:Reference number: AS3420; MUID:95011581; PMID:7926768
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <STR>

A;Residues: 522-917 <CHE>
A;Cross-references: UNIPARC:UPI000016DB96; EMBL:X13435; NID:G16403; PIDN:CAA31787.1; PII:
A;Note: the translation of the nucleotide sequence is not complete
C;Comment: This enzyme catalyzes the reduction of nitrate to nitrite in cytoplasm; each
me involved in the first step of nitrate assimilation in plants, fungi, and bacteria.
C;Genetics:
A;Gene: NIA2
A;Map position: 1
C;Complex: homodimer
C;Superfamily: nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reduc
C;Keywords: chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; ni
F;87-482/Domain: molybdopterin-binding domain homology <PCO>
F;542-616/Domain: cytochrome b5 core homology <CB5>
F;667-917/Domain: cytochrome-b5 reductase homology <CBR>
F;191/Binding site: molybdopterin (Cys) (covalent) #status predicted
F;433/Dissulfide bonds: interchain #status predicted
F;577,600/Binding site: heme iron (His) (axial ligands) #status predicted
F;711,889/Binding site: NAD (Lys, Cys) #status predicted
F;711/Binding site: FAD (Tyr) #status predicted

Query Match 51.3%; Score 39; DB 1; Length 917;
Best Local Similarity 35.7%; Pred. NO. 85;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSEN 14
Db 47 VFLKPAKVHDDDED 60

RESULT 9
B69387
fumarase (fum-2) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: B69387
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
F; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Arttich, P.; Kaine, B.P.; Sykes, S.F.
Smith, H.O.; Weese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69387
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-272 <KLE>
A;Cross-references: UNIPROT:O29166; UNIPARC:UPI0000056E7B; GB:AE001028; GB:AE000782; NID
C;Superfamily: iron-dependent tartrate dehydratase alpha chain; iron-dependent tartrate
F;51-243/Domain: iron-dependent tartrate dehydratase alpha chain homology <TTDA>

Query Match 50.7%; Score 38.5; DB 2; Length 272;
Best Local Similarity 50.0%; Pred. NO. 29;
Matches 10; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

Qy 1 IVLPRSRVH-----NSEENT 15
Db 100 IPLRPNAVHPITRENSGDNT 119

RESULT 10
D71571
hypothetical protein CT326.1 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 30-Oct-1998 #sequence_revision 30-Oct-1998 #text_change 09-Jul-2004
C;Accession: D71571
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: D71571
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-62 <ARN>
A;Cross-references: UNIPROT:O84329; UNIPARC:UPI00000C0B30; GB:AB001305; GB:AB001273; NID
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: CT326.1

Query Match 50.0%; Score 38; DB 2; Length 62;
Best Local Similarity 66.7%; Pred. No. 7.4;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 RVHNSSENT 15
: : : : :
Db 16 RHNSDEENT 24

RESULT 11

S74287
hypothetical protein YCL010c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74287; S19337
R;Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, September 1996
A;Reference number: S74277
A;Accession: S74287
A;Molecule type: DNA
A;Residues: 1-259 <VO>
A;Cross-references: UNIPROT:P25554; UNIPARC:UPI0000053131; EMBL:X59720; NID:G1907116; PI
R;Oliver, S.G.; Anwar, R.; Brown, A.; Gent, M.E.; Indge, K.J.; James, C.M.; Stateva, L.I
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19337
A;Accession: S19337
A;Molecule type: DNA
A;Residues: 1-140, 'TANDTV' <OLI>
A;Cross-references: UNIPARC:UPI000017CC9D; EMBL:X59720; MIPS:YCL010c
A;Note: this sequence has been revised in reference S74287
C;Genetics:
A;Cross-references: SGD:S0000516
A;Map position: 3L
A;Note: YCL010c

Query Match 50.0%; Score 38; DB 2; Length 259;
Best Local Similarity 53.3%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSENT 15
: : : : :
Db 73 LMLNQSREKSEENT 87

RESULT 12

S84487
GTP-binding protein DRG homolog - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G7112; hypothetical protein YGR173w
C;Species: Saccharomyces cerevisiae
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: S64487
R;Hebling, U.; Hofmann, B.; Dellus, H.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64003
A;Accession: S64487
A;Molecule type: DNA
A;Residues: 1-368 <HEB>
A;Cross-references: UNIPROT:P53295; UNIPARC:UPI000013AF8C; EMBL:Z72958; NID:G1323305; PI
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YGR173w
A;Cross-references: SGD:S0003405
A;Map position: 7R
C;Superfamily: GTP-binding protein DRG; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;64-183/Domain: translation elongation factor Tu homology <ETU>
F;70-77/Region: nucleotide-binding motif A (P-loop)

F;93-98/Region: GTP binding #status predicted
F;116-119/Region: GTP binding #status predicted
F;250-253/Region: GTP binding #status predicted
F;341-345/Region: GTP binding #status predicted

Query Match 50.0%; Score 38; DB 1; Length 368;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNS 12
: : : : :
Db 207 MILRDYRIHNAE 218

RESULT 13

T04192
hypothetical protein T4F9.30 - Arabidopsis thaliana
N;Alternate names: protein T4F9.30
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04192
R;Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15260
A;Accession: T04192
A;Molecule type: DNA
A;Residues: 1-328 <BEV>
A;Cross-references: UNIPROT:Q9T0B6; UNIPARC:UPI00000A4AD0; EMBL:AL049523
A;Experimental source: cultivar Columbia; BAC clone T4F9
C;Genetics:
A;Map position: 4
A;introns: 51/2; 122/2; 161/3; 181/3; 214/3; 328/1; 401/3; 458/3; 544/3; 596/2; 774/2; 8
A;Note: T4F9.30
C;Superfamily: ubiquitin C-terminal hydrolase Ubpl2

Query Match 50.0%; Score 38; DB 2; Length 928;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PSRVHNSSEN 14
: : : : :
Db 645 PSKHNGSDN 654

RESULT 14

T37816
probable peroxin-6, AAA family atpase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37816
R;Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21737
A;Accession: T37816
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-948 <DEV>
A;Cross-references: UNIPROT:O13764; UNIPARC:UPI00001316EF; EMBL:Z98849; PIDN:CAB11501.1
A;Experimental source: strain 972h-; cosmid c17A5
C;Genetics:
A;Gene: SPDB:SPAC17A5.01
A;Map position: 1
A;introns: 452/3

Query Match 50.0%; Score 38; DB 2; Length 948;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLPRSRVH 9
: : : : :
Db 181 LVIRPSKVH 189

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHNSSENT 15
|:|:|:|:
Db 52 PPRLHNSQVNS 62

RESULT 20
T23403
hypothetical protein K07C5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C:Accession: T23403
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19737
A:Accession: T23403
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-301 <WIL>
A:Cross-references: UNIPROT:Q21274; UNIPARC:UPI0000079B5F; EMBL:Z71181; PIDN:CAA94895.1;
A:Experimental source: clone K07C5
C:Genetics:
A:Gene: CESP:K07C5.2
A:Map position: 5
A:Introns: 68/3; 112/2; 152/1; 268/3
C:Superfamily: Aldehyde reductase

Query Match 48.7%; Score 37; DB 2; Length 301;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHNSSENT 15
|:|:|:|:
Db 259 PSRIHNDLHNT 269

RESULT 21
B43649
avirulence protein C - Pseudomonas syringae pv. glycinea
C:Species: Pseudomonas syringae pv. Glycinea
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 05-Oct-2004
C:Accession: B43649
R:Tanaki, S.; Dahlbeck, D.; Staekawicz, B.; Keen, N.T.
J. Bacteriol. 170: 4846-4854, 1988
A:Title: Characterization and expression of two avirulence genes cloned from Pseudomonas
A:Reference number: A43649; MUID:89008107; PMID:3049552
A:Accession: B43649
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <TAM>
A:Cross-references: UNIPROT:P13036; UNIPARC:UPI0000126683; GB:M22219; NID:gl51052; PIDN:
C:Superfamily: avirulence protein

Query Match 48.7%; Score 37; DB 2; Length 352;
Best Local Similarity 53.8%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IIVLRPSRVHNSSE 13
:|:|:|:|:
Db 4 VCFRPSRSHVSQE 16

RESULT 22
C81960
hypothetical protein NMA0432 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: C81960
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
R.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: C81960

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <PAR>

A:Cross-references: UNIPROT:Q9QWD3; UNIPARC:UPI00000C4986; GB:ALI62753; GB:ALI57959; NID

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0432

C:Superfamily: Neisseria meningitidis hypothetical protein NMA0432

Query Match 48.7%; Score 37; DB 2; Length 375;

Best Local Similarity 54.5%; Pred. No. 77;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSENT 14

|:|:|:|:
Db 167 RPCRHLNLRQN 177

RESULT 23

T32994

hypothetical protein W02F12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T32994

R:Du, Z.; Gattung, S.; Andrews, S.

submitted to the EMBL Data Library, February 1998

A:Description: The sequence of C. elegans cosmid W02F12.

A:Reference number: Z21261

A:Accession: T32994

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-388 <DUZ>

A:Cross-references: UNIPROT:O45146; UNIPARC:UPI0000079B38; EMBL:AF047653; PIDN:AAC04459

A:Experimental source: strain Bristol N2; clone W02F12

C:Genetics:

A:Gene: CESP:W02F12.3

A:Map position: 5

A:Introns: 27/2; 118/3; 345/3

Query Match 48.7%; Score 37; DB 2; Length 388;

Best Local Similarity 70.0%; Pred. No. 79;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 SRVHNSSENT 15

|:|:|:|:
Db 106 NRVSNTTEENT 115

RESULT 24

A46226

somatostatin receptor 3 - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A46226; S32501

R:Yanada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; S.

Mol. Endocrinol. 6: 2136-2142, 1992

A:Title: Somatostatin receptors, an expanding gene family: cloning and functional charac

A:Reference number: A46226; MUID:93149123; PMID:1337145

A:Accession: A46226

A:Molecule type: DNA

A:Residues: 1-418 <YAM>

A:Cross-references: UNIPROT:P32745; UNIPARC:UPI0000050435; GB:M96738; NID:g338498; PIDN:

A:Note: sequence extracted from NCBI backbone (NCBI:123685, NCBI:123690)

R:Corne, J.D.; Demchisyn, L.L.; Seeman, P.; van Tol, H.H.M.; Srikant, C.B.; Kent, G.

FEBS Lett. 321, 279-284, 1993

A:Title: A human somatostatin receptor (SSTR3), located on chromosome 22, displays pref

A:Reference number: S32501; MUID:93238970; PMID:8097479

A:Accession: S32501

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-418 <COR>

A;Cross-references: UNIPARC:UPI0000050435
C;Genetics:
A;Gene: GDB:SSTR3
A;Cross-references: GDB:I34187; OMIM:182453
F;81-106/Domain: transmembrane #status predicted <TM1>
A;Map position: 22q13.1-22q13.1
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
F;44-70/Domain: transmembrane #status predicted <TM1>
F;118-106/Domain: transmembrane #status predicted <TM2>
F;118-139/Domain: transmembrane #status predicted <TM3>
F;159-181/Domain: transmembrane #status predicted <TM4>
F;203-233/Domain: transmembrane #status predicted <TM5>
F;255-282/Domain: transmembrane #status predicted <TM6>
F;289-316/Domain: transmembrane #status predicted <TM7>
F;17-30/Binding site: carbohydrate (Aen) (covalent) #status predicted
F;116-191/disulfide bonds: #status predicted
F;151-251-317-332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;231/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predicted
F;236/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predicted
F;412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 48.7%; Score 37; DB 2; Length 418;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEENT 15
:|||||:|:
Db 327 VLLRPSRVRSQPT 341

RESULT 25
H84379
4-hydroxybutyrate CoA transferase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84379
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitthauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcig, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Oner, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: H84379
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-491 <STO>
A;Cross-references: UNIPROT:Q9HNL7; UNIPARC:UPI0000063ACE; GB:AE004437; NID:gl0581703; P
C;Genetics:
A;Gene: cat
C;Superfamily: acetyl-CoA hydrolase

Query Match 48.7%; Score 37; DB 2; Length 491;
Best Local Similarity 63.6%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNS 11
:|||||:|:
Db 338 VLLRPSRVNSNA 348

RESULT 26
G65076
hypothetical protein b2920 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: G65076
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: G65076
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-492 <BLAT>
A;Cross-references: UNIPROT:P52043; UNIPARC:UPI000013B07C; GB:AE000375; GB:U00096; NID:gl
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: acetyl-CoA hydrolase

Query Match 48.7%; Score 37; DB 2; Length 492;
Best Local Similarity 53.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSE 13
:|||||:|:
Db 330 IVLRPQEIENNPE 342

RESULT 27
F91102
probable coenzyme A transferase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F91102
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <HAY>
A;Cross-references: UNIPROT:Q8XD13; UNIPARC:UPI00000D03EF; GB:BA000007; PIDN:BAB37213.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3790
C;Superfamily: acetyl-CoA hydrolase

Query Match 48.7%; Score 37; DB 2; Length 492;
Best Local Similarity 53.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSE 13
:|||||:|:
Db 330 IVLRPQEIENNPE 342

RESULT 28
B85948
probable coenzyme A transferase ygfH [imported] - Escherichia coli (strain O157:H7, subs
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85948
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85948
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <STO>
A;Cross-references: UNIPROT:Q8XD13; UNIPARC:UPI00000D03EF; GB:AE005174; NID:gl12517452; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ygfH
C;Superfamily: acetyl-CoA hydrolase

Query Match 48.7%; Score 37; DB 2; Length 492;
Best Local Similarity 53.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSE 13

```
||||| : | : |
Db 330 IVLRPQEISNPN 342

RESULT 29
A47726
dis1-suppressing protein kinase dsk1 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004
C:Accession: A47726; T40530
R:Takeuchi, M.; Yanagida, M.
Mol. Biol. Cell 4, 247-260, 1993
A:Title: A mitotic role for a novel fission yeast protein kinase dsk1 with cell cycle sb
A:Reference number: A47726; MUID:93250320; PMID:8485317
A:Accession: A47726
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-544 <TAK>
A:Cross-references: UNIPROT:E36616; UNIPARC:UPI0000144718; GB:D13447; NID:G931945; PIDN:
A>Note: sequences extracted from NCBI backbone (NCBIN:131089, NCBI:P:131111)
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21934
A:Accession: T40530
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-40, 'A' 42-544 <LYN>
A:Cross-references: UNIPARC:UPI0000129903; EMBL:AL023634; PIDN:CAA19180.1; GSPDB:GN00067
A:Experimental source: strain 972h-; cosmid c530
C:Genetics:
A:Map position: 2
A:Gene: SPDB:SPBC530.14C
A:Superfamily: protein kinase, CLK type; protein kinase homology
C:Keywords: serine/threonine-specific protein kinase

Query Match 48.7%; Score 37; DB 2; Length 544;
Best Local Similarity 53.3%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 15
| | | | |
Db 316 ISLRDSQKSHSPNS 330

RESULT 30
AC3547
ribose transport ATP-binding protein rbsA BMEI10300 [imported] - Brucella melitensis (st
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AC3547
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3547
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-584 <KUR>
A:Cross-references: UNIPROT:OBYD77; UNIPARC:UPI00005841B; GB:AE008918; PIDN:AAL53542.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10300
A:Map position: 11
C:Superfamily: Bacillus subtilis probable ABC transporter yufO; ATP-binding cassette hom

Query Match 48.7%; Score 37; DB 2; Length 584;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVH 9
| | | | |
Db 110 IALRPGRVH 118
```

RESULT 31

B84902

hypothetical protein At2g46380 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: B84902

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84902

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-732 <STO>

A:Cross-references: UNIPROT:Q9SKE1; UNIPARC:UPI000000AB79C; GB:AE002093; NID:G4559381; PI

C:Genetics:

A:Gene: At2g46380

A:Map position: 2

Query Match 48.7%; Score 37; DB 2; Length 732;

Best Local Similarity 54.5%; Pred. No. 1.6e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13

| | | | |

Db 248 LQPSHLNSRQ 258

RESULT 32

AC0403

catalase (EC 1.11.1.6) [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004

C:Accession: AC0403

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0403

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-737 <KUR>

A:Cross-references: UNIPROT:Q9X6B0; UNIPARC:UPI0000127086; GB:AL590842; PIDN:CAC92551.1;

C:Genetics:

A:Gene: katY

C:Superfamily: Catalase HPI

C:Keywords: oxidoreductase

Query Match 48.7%; Score 37; DB 2; Length 737;

Best Local Similarity 58.3%; Pred. No. 1.6e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 14

| | | | |

Db 40 LSPLRLHNIESN 51

RESULT 33

T31994

hypothetical protein C49D10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T31994

R:Henkhaus, J.; Wohlmann, P.; Beck, C.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid C49D10.

A:Reference number: Z21108

A;Accession: T31994
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-758 <HEN>
A;Cross-references: UNIPROT:O16610; UNIPARC:UPI000007F7DE; EMBL:AF016665; PIDN:AAC71186.
A;Experimental source: strain Bristol N2; clone C49D10
C;Genetics:
A;Gene: CESP:C49D10.1
A;Map position: 2
A;Introns: 438/2

Query Match 48.7%; Score 37; DB 2; Length 758;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNS 11
:|:|:|:|:|
DB 57 MVTRPSRISNS 67
:|:|:|:|:|

RESULT 34
T04982
dynam-in-like protein ADL2 - Arabidopsis thaliana
N;Alternate names: protein T16L1.140
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
A;Accession: T04982
R;Bevan, M.; Obermaier, B.; Deutschenbaur, S.; Piravandi, E.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, November 1998
A;Reference number: Z15393
A;Accession: T04982
A;Molecule type: DNA
A;Residues: 1-808 <BEV>
A;Cross-references: UNIPROT:O81982; UNIPARC:UPI000009DDF4; EMBL:AL031394
A;Experimental source: cultivar Columbia; BAC clone T16L1
C;Genetics:
A;Gene: ADL2
A;Map position: 4
A;Introns: 138/3; 233/1; 242/3; 276/3; 299/3; 319/3; 357/3; 374/1; 410/3; 441/3; 472/3;
A;Note: T16L1.140
C;Superfamily: dynam-in-related protein VP81
C;Keywords: P-loop; purine nucleotide binding
F;66-73/Region: nucleotide-binding motif A (P-loop)
F;168-173/Region: nucleotide-binding motif B

Query Match 48.7%; Score 37; DB 1; Length 808;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLRPSRVHNS 13
:|:|:|:|:|
DB 655 VLRPTETHSEQE 666
:|:|:|:|:|

RESULT 35
D64119
glycogen phosphorylase (EC 2.4.1.1) - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
A;Accession: D64119
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A.; Title: Whole-Genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: D64119
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-821 <TIGR>
A;Cross-references: UNIPROT:P45180; UNIPARC:UPI0000131A14; GB:U32815; GB:L42023; NID:915
C;Superfamily: Glucan phosphorylase

C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphate
F;567/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 48.7%; Score 37; DB 2; Length 821;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNS 14
:|:|:|:|:|
DB 302 ILRRHKRTHNSLEN 315
:|:|:|:|:|

RESULT 36
T14470
receptor-like kinase (EC 2.7.1.1-) SPR2 - wild cabbage
C;Species: Brassica oleracea (wild cabbage)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 31-Dec-2004
A;Accession: T14470; T14445
R;Pastuglia, M.; Roby, D.; Dumas, C.; Cock, J.M.
Plant Cell 9, 1-13, 1997
A;Title: Rapid induction by wounding and bacterial infection of an S gene family receptor
A;Reference number: Z18107
A;Accession: T14470
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-847 <PAS>
A;Cross-references: UNIPROT:P93068; UNIPARC:UPI00000A1551; EMBL:X98520
A;Experimental source: variety acephala, cultivar P57SI
R;Stanchev, B.S.; Croy, R.R.D.
submitted to the EMBL Data Library, April 1997
A;Reference number: Z18094
A;Accession: T14445
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-23, 'S', 25-622, 'V', 624-847 <STA>
A;Cross-references: UNIPARC:UPI00000A2F9E; EMBL:Y12530
A;Experimental source: strain S29
C;Genetics:
A;Gene: SPR2; ARUK
A;Introns: 429/1; 474/1; 538/3; 609/1; 688/2; 738/3
C;Superfamily: S-locus receptor-like kinase SRK; protein kinase homology; S-locus-specific
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tra
F;32-427/Domain: S-locus-specific glycoprotein homology <SSG>
F;520-804/Domain: protein kinase homology <KIN>

Query Match 48.7%; Score 37; DB 2; Length 847;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNS 14
:|:|:|:|:|
DB 487 WVPPNRRHSREN 500
:|:|:|:|:|

RESULT 37
S74756
hypothetical protein slr1613 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
A;Accession: S74756
R;Kaneke, T.; Sato, S.; Kofani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74756
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-196 <KAN>
A;Cross-references: UNIPROT:P72891; UNIPARC:UPI00000C0C38; EMBL:D90901; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: E97975
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <KUR>
A;Cross-references: UNIPROT:Q8DQ63; UNIPARC:UPI00000E3526; GB:AE007317; PIDN:AAK99634.1;
C;Genetics:
A;Gene: rlpD
C;Superfamily: conserved hypothetical protein HI0176

Query Match 47.4%; Score 36; DB 2; Length 295;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSENT 15
||||| : : :
Db 122 VLPRGIVHRIDKDT 135

RESULT 43
E97160
probable pseudouridylylate synthase, YLYB B. subtilis ortholog [imported] - Clostridium ac
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97160
R;Nolling, J.; Brston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97160
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <KUR>
A;Cross-references: UNIPROT:Q97H99; UNIPARC:UPI00000CA400; GB:AE001437; PIDN:AAK90072.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2114
C;Superfamily: conserved hypothetical protein HI0176

Query Match 47.4%; Score 36; DB 2; Length 305;
Best Local Similarity 50.0%; Pred. No. 94;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSENT 15
||||| : : :
Db 128 VLPRGIVHRIDKDT 141

RESULT 44
DTECC
aspartate carbamoyltransferase (EC 2.1.3.2) catalytic chain - Escherichia coli (strain K
N;Alternate names: aspartate transcarbamylase catalytic chain; aspartyl carbamoyltransfe
C;Species: Escherichia coli
C;Date: 30-Nov-1980 #sequence_revision 10-Oct-1997 #text_change 31-Dec-2004
C;Accession: H65236; A00561; A21121; E36599; A21120; S56471
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H65236
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <BLAT>
A;Cross-references: UNIPROT:P00479; UNIPARC:UPI0000111347; GB:AE0000495; GB:U00096; MID:9
A;Experimental source: strain K-12, substrain MG1655
R;Schachman, H.K.; Pauza, C.D.; Navre, M.; Karels, M.J.; Wu, L.; Yang, Y.R.
Proc. Natl. Acad. Sci. U.S.A. 81, 115-119, 1984
A;Title: Location of amino acid alterations in mutants of aspartate transcarbamoylase: s
A;Reference number: A93985; MUID:84119419; PMID:6364131
A;Accession: A00561
A;Molecule type: DNA

A;Residues: 1-149,'E',151-311 <SCH>
A;Cross-references: UNIPARC:UPI000016F43B; GB:K01472; MID:g147463; PIDN:AAA24476.1; PID:9
R;Hoover, T.A.; Roof, W.D.; Foltermann, K.F.; O'Donovan, G.A.; Bencini, D.A.; Wild, J.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 2462-2466, 1983
A;Title: Nucleotide sequence of the structural gene (pyrB) that encodes the catalytic pol
A;Reference number: A21121; MUID:83195078; PMID:6302686
A;Accession: A21121
A;Molecule type: DNA
A;Residues: 1-60,'O',62-165,'T',167-220,'V',222-311 <HOO>
A;Cross-references: UNIPARC:UPI0000172336; GB:V00323
R;Donahue, J.P.; Turnbough Jr., C.L.
J. Biol. Chem. 265, 19091-19099, 1990
A;Title: Characterization of transcriptional initiation from promoters P-1 and P-2 of the
A;Reference number: A36599; MUID:91035438; PMID:1699940
A;Accession: B36599
A;Molecule type: DNA
A;Residues: 1-18 <DON>
A;Cross-references: UNIPARC:UPI000016F43D; GB:M60508; MID:g147469; PIDN:AAA24481.1; PID:9
R;Konigsberg, W.H.; Henderson, L.
Proc. Natl. Acad. Sci. U.S.A. 80, 2467-2471, 1983
A;Title: Amino acid sequence of the catalytic subunit of aspartate transcarbamoylase from
A;Reference number: A21120; MUID:83195079; PMID:6341995
A;Accession: A21120
A;Molecule type: protein
A;Residues: 2-60,'Q',62-86,'Q',88-90,'N',92-129,'N',131-220,'V',222-256,'D',258-259,'M',2
A;Cross-references: UNIPARC:UPI0000172337
R;Ye, H.; Honzato, R.B.; Lipscomb, W.N.
Proc. Natl. Acad. Sci. U.S.A. 81, 4037-4040, 1984
A;Title: Structure of unligated aspartate carbamoyltransferase of Escherichia coli at 2.6
A;Reference number: A93993; MUID:84248054; PMID:6377306
A;Contents: annotation; X-ray crystallography, 2.6 angstroms; quaternary structure
R;Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A;Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.6
A;Reference number: S56314; MUID:95334362; PMID:7610040
A;Accession: S56471
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-195,'R',197-311 <BUR>
A;Cross-references: UNIPARC:UPI000016F6D0; EMBL:U14003; MID:g1263172; PIDN:AAA97142.1; PY
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C;Comment: The active enzyme contains two trimers of catalytic chains and three dimers of
C;Genetics:
A;Gene: pyrB
A;Map position: 97 min
C;Superfamily: ornithine/aspartate carbamoyltransferase; aspartate/ornithine carbamoyltra
C;Keywords: heterododecamer; homohexamer; homotrimer; pyrimidine nucleotide biosynthesis
F;8-305/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match 47.4%; Score 36; DB 1; Length 311;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSENT 14
||||| : : :
Db 249 VLPRASDLHNKAN 261

RESULT 45
G86122
hypothetical protein pyrB [imported] - Escherichia coli (strain O157:H7, substrain EDL93:
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C;Accession: G86122
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G86122
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <STO>

A;Cross-references: UNIPARC:UPI00001659FA; GB:AE005174; NID:gl2519253; PIDN:AAGS9443.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: pyrB
C;Superfamily: ornithine/aspartate carbamoyltransferase; aspartate/ornithine carbamoyltransferase

Query Match 47.4%; Score 36; DB 2; Length 311;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
|||||:|
Db 249 VLRAADLHNAKAN 261

RESULT 46
F91281
aspartate carbamoyltransferase (EC 2.1.3.2) catalytic chain - Escherichia coli (strain C)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 31-Dec-2004
C;Accession: F91281
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic analysis
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91281
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <HAY>
A;Cross-references: UNIPROT:P00479; UNIPARC:UPI0000111347; GB:BA0000007; PIDN:BA838645.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC85222
C;Superfamily: ornithine/aspartate carbamoyltransferase; aspartate/ornithine carbamoyltransferase

Query Match 47.4%; Score 36; DB 2; Length 311;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
|||||:|
Db 249 VLRAADLHNAKAN 261

RESULT 47
T34873
hypothetical protein SC3A3.08 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Oct-2004
C;Accession: T34873
R;Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21560
A;Accession: T34873
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-392 <SEE>
A;Cross-references: UNIPROT:Q9S2Y5; UNIPARC:UPI00000DB332; EMBL:AL109849; PIDN:CAB52865.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC0EDB:SC3A3.08
C;Superfamily: ADP-ribosylglycohydrolase

Query Match 47.4%; Score 36; DB 2; Length 392;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSE 13
|||||:|
Db 281 RPSRLHAIEE 290

RESULT 48
S44778
transcription factor unc-86, short splice form - Caenorhabditis elegans
N;Alternate names: C30A5.6 protein
C;Species: Caenorhabditis elegans
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S44778
R;Anderson, K.
submitted to the EMBL Data Library, February 1993
A;Description: Sequence of the C. elegans cosmid C30A5.
A;Reference number: S44774
A;Accession: S44778
A;Molecule type: DNA
A;Residues: 1-429 <AND>
A;Cross-references: UNIPROT:P13528; UNIPARC:UPI000002B694; EMBL:L10990; NID:gl56211; PIDN:
C;Genetics:
A;Introns: 128/3; 253/3; 345/2; 382/2
C;Superfamily: transcription factor unc-86; homeobox homology; POU domain homology
C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
F;234-304/Domain: POU domain homology <POU>
F;326-382/Domain: homeobox homology <HOX>

Query Match 47.4%; Score 36; DB 1; Length 429;
Best Local Similarity 46.2%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
|||||:|
Db 46 LLRPSKISRGSEN 58

RESULT 49
A30042
transcription factor unc-86, long splice form - Caenorhabditis elegans
N;Alternate names: C30A5.7 protein
C;Species: Caenorhabditis elegans
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: A30042; S44779
R;Finney, M.; Ruvkun, G.; Horvitz, H.R.
Cell 55, 757-769, 1988
A;Title: The C. elegans cell lineage and differentiation gene unc-86 encodes a protein w
A;Reference number: A30042; MUID:89051866; PMID:2903797
A;Accession: A30042
A;Molecule type: DNA
A;Residues: 1-467 <FIN>
A;Cross-references: UNIPROT:P13528; UNIPARC:UPI0000137C53; GB:L10990; NID:gl56211; PIDN:
R;Anderson, K.
submitted to the EMBL Data Library, February 1993
A;Description: Sequence of the C. elegans cosmid C30A5.
A;Reference number: S44774
A;Accession: S44779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-467 <AND>
A;Cross-references: UNIPARC:UPI0000137C53; EMBL:L10990; NID:gl56211; PIDN:AAB59176.1; PI
C;Genetics:
A;Gene: unc-86
A;Introns: 128/3; 174/2; 291/3; 383/2; 420/2
C;Superfamily: transcription factor unc-86; homeobox homology; POU domain homology
C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
F;272-342/Domain: POU domain homology <POU>
F;364-420/Domain: homeobox homology <HOX>

Query Match 47.4%; Score 36; DB 1; Length 467;
Best Local Similarity 46.2%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
|||||:|
Db 46 LLRPSKISRGSEN 58

RESULT 50

H97289
dihydroxy-acid dehydratase [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 05-Oct-2004
C/Accession: H97289
R/Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld

A/Reference number: A96900; PMID:21359325; PMID:21359325

A/Accession: H97289

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-552 <KUR>

A/Cross-references: UNIPROT:Q97EE3; UNIPARC:UPI000012D573; GB:AE001437; PIDN:AAK81107.1;

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Gene: CAC3170

C/Superfamily: Dihydroxy-acid dehydratase/phosphogluconate dehydratase

Query Match 47.4%; Score 36; DB 2; Length 552;

Best Local Similarity 77.8%; Pred. No. 1.8e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PSRVHSEE 13

Db 401 PARVFNSEE 409

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

A/Accession: T31163

A/Status: preliminary; translated from GB/EMBL/DBSJ

A/Molecule type: DNA

A/Residues: 1-595 <ROM>

A/Cross-references: UNIPROT:O85872; UNIPARC:UPI000005C97C; EMBL:AF079317; NID:G33378261; I

C/Genetics:

A/Genome: plasmid pNL1

A/Note: orf347

Query Match 47.4%; Score 36; DB 2; Length 595;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 SRVHSEEN 14

Db 341 ARVHSSDN 349

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 51

G75622
probable oligopeptide transport periplasmic protein - Deinococcus radiodurans (strain R1
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: G75622
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Otterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; PMID:20036896; PMID:10567266

A/Accession: G75622

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-578 <WHI>

A/Cross-references: UNIPROT:Q9R2S2; UNIPARC:UPI00000D3B45; GB:AE001826; NID:G6460827; PI

A/Experimental source: strain R1

C/Genetics:

A/Gene: DRB0042

A/Map position: megaplasmid

A/Genome: plasmid

A/Note: plasmid MP1

Query Match 47.4%; Score 36; DB 2; Length 578;

Best Local Similarity 66.7%; Pred. No. 1.8e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHSE 12

Db 397 RPARVHSE 405

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 52

T31163
hypothetical protein orf347 - Sphingomonas aromaticivorans plasmid pNL1
C/Species: Sphingomonas aromaticivorans
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T31163
R/Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Senses, C.W.; G
submitted to the EMBL Data Library, July 1998

A/Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati

A/Reference number: Z20992

A/Accession: T31163

A/Status: preliminary; translated from GB/EMBL/DBSJ

A/Molecule type: DNA

A/Residues: 1-634 <WIL>

A/Cross-references: UNIPROT:O46020; UNIPARC:UPI0000082CA1; EMBL:Z92822; PIDN:CAB07301.1;

A/Experimental source: clone ZK520

C/Genetics:

A/Gene: CESP:ZK520.3

A/Map position: 3

A/Introns: 53/2; 75/2; 120/2; 187/1; 318/2; 341/1; 455/1; 530/3

Query Match 47.4%; Score 36; DB 2; Length 634;

Best Local Similarity 58.3%; Pred. No. 2e+02;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRPSRVHSEEN 14

Db 3 LRPSRVHSEEN 14

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Db 490 LRLHRINSEDN 501

RESULT 55

T34573

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C:Accession: T34573

R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z21548

A:Accession: T34573

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-691 <MUR>

A:Cross-references: UNIPROT:O54098; UNIPARC:UPI00000DABC8; EMBL:AL021529; PIDN:CAA16437.

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SC0EDB:SC10A5.05

Query Match 47.4%; Score 36; DB 2; Length 691;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLPRSRVHN 10

:::|||||:

Db 512 WVRPSRVHS 520

RESULT 56

S56277

probable membrane protein YFR022w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004

C:Accession: S56277

R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanuma

submitted to the EMBL Data Library, May 1995

A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce

A:Reference number: S56186

A:Accession: S56277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-733 <MUR>

A:Cross-references: UNIPROT:P43602; UNIPARC:UPI0000036C46; EMBL:D50617; NID:G836685; PID

C:Genetics:

A:Cross-references: SGD:S0001918

A:Map position: 6R

C:Superfamily: Saccharomyces cerevisiae probable membrane protein YFR022w

C:Keywords: transmembrane protein

F:382-398/Domain: transmembrane #status predicted <TM>

Query Match 47.4%; Score 36; DB 2; Length 733;

Best Local Similarity 50.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHNSNT 15

:::|||||:

Db 639 KPSRVHNSNT 650

RESULT 57

T47503

hypothetical protein F9K21.210 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T47503

R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24467

A:Accession: T47503

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-989 <JOR>

A:Cross-references: UNIPROT:Q9M1E4; UNIPARC:UPI00000A8243; EMBL:AL138657

A:Experimental source: cultivar Columbia; BAC clone F9K21

C:Genetics:

A:Map position: 3

A:Introns: 31/3; 74/2; 124/3; 236/2; 205/3; 189/2; 161/2; 124/3; 305/2; 278/3; 757/1; 794/1

A:Note: F9K21.210

Query Match 47.4%; Score 36; DB 2; Length 989;

Best Local Similarity 46.7%; Pred. No. 3.3e+02;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSNT 15

:::|||||:

Db 418 VIDKPSAVNSLEHT 432

RESULT 58

T45283

growth polarity maintenance protein st55 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004

C:Accession: T45283

R:Toda, T.; Niwa, H.; Nemoto, T.; Dhut, S.; Eddison, M.; Matsusaka, T.; Yanagida, M.; Hi

J. Cell Sci. 109, 2331-2342, 1996

A:Title: The fission yeast st55+ gene is required for maintenance of growth polarity and

A:Reference number: Z22956; MUID:97041712; PMID:8886983

A:Accession: T45283

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1066 <TOD>

A:Cross-references: UNIPROT:O74454; UNIPARC:UPI00001688DC; EMBL:D58421; NID:G2645172; PI

C:Genetics:

A:Gene: st55

C:Keywords: signal transduction

Query Match 47.4%; Score 36; DB 2; Length 1066;

Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSNT 15

:::|||||:

Db 450 ILRPSAANKERQT 463

RESULT 59

T41099

staurosporine target st5p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T41099

R:Purnelle, B.; Goffeau, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21964

A:Accession: T41099

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1066 <PUR>

A:Cross-references: UNIPROT:O74454; UNIPARC:UPI0000136150; EMBL:AL031535; PIDN:CAA20748.

A:Experimental source: strain 972h-; cosmid c16C4

C:Genetics:

A:Gene: SPDB:SPCC16C4.09

A:Map position: 3

Query Match 47.4%; Score 36; DB 2; Length 1066;

Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSNT 15

:::|||||:

Db 450 ILRPSAANKERQT 463

RESULT 60
 A39578
 SSD1 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D9819.4; protein YDR293c; suppressor protein SRK1
 C:Species: Saccharomyces cerevisiae
 C:Date: 07-Feb-1992 #sequence revision 07-Feb-1992 #text_change 09-Jul-2004
 C:Accession: A39578; A40263; S70122; S14774; S16805
 R:Sutton, A.; Immanuel, D.; Arndt, K.T.
 Mol. Cell. Biol. 11, 2133-2148, 1991
 A:Title: The SIT4 protein phosphatase functions in late G(1) for progression into S phase
 A:Reference number: S14773; MUID:91172202; PMID:1848673
 A:Accession: A39578
 A:Molecule type: DNA
 A:Residues: 1-1250 <SUT>
 A:Cross-references: UNIPROT:P24276; UNIPARC:UPI0000135FBA; GB:M60318; NID:gl72611; PIDN:
 R:Wilson, R.B.; Brenner, A.A.; White, T.B.; Engler, M.J.; Gaughran, J.P.; Tatchell, K.
 Mol. Cell. Biol. 11, 3369-3373, 1991
 A:Title: The Saccharomyces cerevisiae SRK1 gene, a suppressor of bcy1 and ins1, may be
 A:Reference number: A40263; MUID:91246207; PMID:1645449
 A:Accession: A40263
 A:Molecule type: DNA
 A:Residues: 1-1250 <WIL>
 A:Cross-references: UNIPARC:UPI0000135FBA; GB:M63004; NID:gl72696; PIDN:AAA35089.1; PID:
 R:Fulton, L.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of S. cerevisiae cosmid 9819.
 A:Reference number: S70114
 A:Accession: S70122
 A:Molecule type: DNA
 A:Residues: 1-1250 <FUL>
 A:Cross-references: UNIPARC:UPI0000135FBA; EMBL:U51031; NID:gl332635; PID:gl230657; MIPS
 C:Genetics:
 A:Gene: SGD:SSD1; SRK1
 A:Cross-references: SGD:S0002701; MIPS:YDR293c
 A:Map position: 4R

Query Match 47.4%; Score 36; DB 2; Length 1250;
 Best Local Similarity 53.8%; Pred. No. 4.2e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSEN 14
 :|||: ||: |
 Db 561 LRFPSQQNSDNN 573

Search completed: March 11, 2006, 12:25:34
 Job time : 37.6842 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:12:17 ; Search time 146.842 Seconds
(without alignments)
72.070 Million cell updates/sec

Title: US-10-774-242a-1

Perfect score: 76

Sequence: 1 IVLPRSVHNEENT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	760	1 SEPR HUMAN	Q12884 homo sapien
2	76	100.0	760	2 Q53TP5 HUMAN	Q53TP5 homo sapien
3	53	69.7	350	2 Q6P7D6 RAT	Q6P7D6 rattus norv
4	53	69.7	761	2 Q8R492 RAT	Q8R492 rattus norv
5	52	68.4	761	1 SEPR MOUSE	P97321 mus musculus
6	46	60.5	325	2 Q73QP6 TREDE	Q73QP6 treponema d
7	45	59.2	277	2 Q4RR26 TETNG	Q4RR26 tetraodon n
8	44	57.9	81	2 Q61JW5 DROME	Q61JW5 drosophila
9	44	57.9	809	2 Q61710 CASER	Q61710 caenorhabdi
10	43	56.6	862	2 Q4WU2 ASPFU	Q4WU2 aspergillus
11	43	56.6	1116	2 Q5AT11 EMENI	Q5AT11 aspergillus
12	42	55.3	257	2 Q4Q3F8 LEIMA	Q4Q3F8 leishmania
13	42	55.3	425	2 Q9NKE8 HUMAN	Q9NKE8 homo sapien
14	42	55.3	435	2 Q6BYR6 DEBHA	Q6BYR6 debaryomyce
15	42	55.3	444	2 Q8YR14 ANASP	Q8YR14 anabaena sp
16	42	55.3	787	2 Q59NX5 CANAL	Q59NX5 candida alb
17	42	55.3	837	2 Q7X125 ORYSA	Q7X125 oryza sativ
18	41	53.9	451	2 Q7NYT4 CHRVO	Q7NYT4 chromobacte
19	41	53.9	499	2 Q5LJ72 BACFN	Q5LJ72 bacteroides
20	41	53.9	499	2 Q65OP9 BACFR	Q65OP9 bacteroides
21	41	53.9	585	2 Q57VF6 YTRYP	Q57VF6 trypanosoma
22	41	53.9	740	2 Q9XUM0 CABEL	Q9XUM0 caenorhabdi
23	41	53.9	817	1 PHSG PRSMU	Q9C9N0 pasteurella
24	41	53.9	887	1 Q9ZSB6 ARATH	Q9ZSB6 arabidopsis
25	41	53.9	910	2 Q93Y01 ARATH	Q93Y01 arabidopsis
26	41	53.9	937	2 Q9T0B8 ARATH	Q9T0B8 arabidopsis
27	41	53.9	2828	2 Q9NR99 HUMAN	Q9NR99 homo sapien
28	40	52.6	115	2 Q8NPC0 CORGL	Q8NPC0 corynebacte
29	40	52.6	143	2 Q5CZC9 BACFN	Q5CZC9 bacteroides
30	40	52.6	174	2 Q8ELX1 OCEIH	Q8ELX1 oceanobacil
31	40	52.6	269	2 Q6TV68_9POXV	Q6TV68 bovine papu

32	40	52.6	310	2 Q6ZRG2 HUMAN	Q6ZRG2 homo sapien
33	40	52.6	350	2 Q6HMI1_BACHK	Q6HMI1 bacillus th
34	40	52.6	350	2 Q81UK7_BACAN	Q81UK7 bacillus an
35	40	52.6	350	2 Q63FB9_BACCZ	Q63FB9 bacillus ce
36	40	52.6	414	2 Q4H676_9DEIO	Q4H676 deinococcus
37	40	52.6	442	1 Q8R4B4 homo sapien	Q8R4B4 homo sapien
38	40	52.6	443	1 Q8R4B4 mouse	Q8R4B4 mus musculu
39	40	52.6	443	1 Q8R4B4 RAT	Q8R4B4 rattus norv
40	40	52.6	491	2 Q6G155_BARQU	Q6G155 bartonella
41	40	52.6	520	2 Q8L775_ARATH	Q8L775 arabidopsis
42	40	52.6	532	1 Q8L775_ARATH	Q8L775 arabidopsis
43	40	52.6	771	2 Q5GZ92_XANOR	Q5GZ92 xanthomonas
44	40	52.6	834	2 Q6TQF0_YANCV	Q6TQF0 xanthomonas
45	40	52.6	836	1 Q8R4B4 mouse	Q8R4B4 mus musculu
46	40	52.6	836	1 Q8R4B4 RAT	Q8R4B4 rattus norv
47	40	52.6	836	1 Q8R4B4 RAT	Q8R4B4 rattus norv
48	40	52.6	836	1 Q8R4B4 RAT	Q8R4B4 rattus norv
49	39	51.3	7716	2 Q8L775_ARATH	Q8L775 arabidopsis
50	39	51.3	113	2 Q8R4B4 mouse	Q8R4B4 mus musculu
51	39	51.3	139	2 Q8R4B4 RAT	Q8R4B4 rattus norv
52	39	51.3	142	1 Q8R4B4 RAT	Q8R4B4 rattus norv
53	39	51.3	142	1 Q8R4B4 RAT	Q8R4B4 rattus norv
54	39	51.3	212	2 Q4R999_MACFA	Q4R999 macaca fasc
55	39	51.3	217	2 Q4RF61_TETNG	Q4RF61 tetraodon n
56	39	51.3	226	2 Q6AL13_DESPS	Q6AL13 desulfotale
57	39	51.3	294	2 Q4P175_USTMA	Q4P175 usilago ma
58	39	51.3	316	2 Q4WQC4_ASPFU	Q4WQC4 aspergillus
59	39	51.3	316	2 Q73G36_WOLPM	Q73G36 wolbachia p
60	39	51.3	333	2 Q9M0V1_ARATH	Q9M0V1 arabidopsis
61	39	51.3	364	2 Q6D519_BRWCT	Q6D519 erwinia car
62	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
63	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
64	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
65	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
66	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
67	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
68	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
69	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
70	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
71	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
72	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
73	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
74	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
75	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
76	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
77	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
78	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
79	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
80	39	51.3	368	2 Q6FW28_CANGA	Q6FW28 candida gla
81	38.5	50.7	272	2 Q29166_ARCFU	Q29166 archaeoglob
82	38.5	50.7	697	2 Q6CMG8_KLULA	Q6CMG8 kluyveromyc
83	38	50.0	62	2 Q84329_CHLTR	Q84329 chlamydia t
84	38	50.0	100	2 Q6YR21_ONYPE	Q6YR21 onion yello
85	38	50.0	104	2 Q7QDL6_ANGOA	Q7QDL6 anopheles g
86	38	50.0	148	2 Q8N9E6_HUMAN	Q8N9E6 homo sapien
87	38	50.0	209	2 Q8AJR8_9HIV1	Q8AJR8 human immun
88	38	50.0	232	2 Q8BVR8_LACPL	Q8BVR8 lactobacill
89	38	50.0	242	2 Q53PL6_ORYSA	Q53PL6 oryza sativ
90	38	50.0	259	1 YCB0_YEAST	YCB0 saccharomyc
91	38	50.0	285	2 Q889C8_PLAFA	Q889C8 plasmodium
92	38	50.0	291	2 Q93TU0_9GAMM	Q93TU0 acidithioba
93	38	50.0	319	2 Q4TFR7_TETNG	Q4TFR7 tetraodon n
94	38	50.0	354	2 Q874Y5_PODAN	Q874Y5 podospora a
95	38	50.0	368	1 YG3Y_YEAST	YG3Y saccharomyc
96	38	50.0	371	2 Q751N0_ASHGO	Q751N0 ashbya gos
97	38	50.0	431	2 Q88Y90_LACPL	Q88Y90 lactobacill
98	38	50.0	490	2 Q62062_ORYSA	Q62062 oryza sativ
99	38	50.0	495	2 Q9VR20_DROME	Q9VR20 drosophila
100	38	50.0	503	2 Q4R190_TETNG	Q4R190 tetraodon n
101	38	50.0	507	2 Q5ARQ0_EMENI	Q5ARQ0 aspergillus
102	38	50.0	569	2 Q6MNM5_BDEBA	Q6MNM5 bdellovibri
103	38	50.0	603	2 Q7FJX4_ANGOA	Q7FJX4 anopheles g
104	38	50.0	696	2 Q4WBM4_ASPFU	Q4WBM4 aspergillus

105	38	50.0	729	2	Q4U9V5	THEAN	Q4u9v5	theileria a	178	37	48.7	428	2	Q4QR01	XENLA	Q4qr01	xenopus lae
106	38	50.0	748	2	Q924H6	RAT	Q924h6	rattus norv	179	37	48.7	432	2	Q7X0V4	STRPY	Q7x0v4	streptococc
107	38	50.0	782	1	OSTA	PASMU	OSTA	pasteurella	180	37	48.7	440	1	STRO	STRPY	STRO	streptococc
108	38	50.0	841	2	Q44207	AGRRH	Q44207	agrobacteri	181	37	48.7	440	2	Q7X0Y6	STRPY	Q7x0y6	streptococc
109	38	50.0	845	2	Q516S9	ENTHI	Q516s9	entamoeba h	182	37	48.7	441	2	Q62Y5	ASCSU	Q62y5	ascaris suu
110	38	50.0	901	2	Q32FG1	MYXXA	Q32fg1	myxococcus	183	37	48.7	457	2	Q51Y51	MAGGR	Q51y51	magnaporthe
111	38	50.0	928	2	Q3T0B6	ARATH	Q3t0b6	arabidopsis	184	37	48.7	462	2	Q9DA23	MOUSE	Q9da23	mus musculus
112	38	50.0	948	1	PEX6	SCHPO	Q13764	echinosacch	185	37	48.7	477	2	Q8G5Z1	BIFLO	Q8g5z1	bifidobacte
113	38	50.0	1028	2	Q9ZS85	ARATH	Q9zsb5	arabidopsis	186	37	48.7	482	2	Q9CVG8	MOUSE	Q9cvg8	mus musculus
114	38	50.0	1278	2	Q962V7	LEIMA	Q96b27	leishmania	187	37	48.7	483	2	Q8KQ01	MOUSE	Q8kq01	mus musculus
115	38	50.0	1371	2	Q96228	PLAF7	Q96228	plasmodium	188	37	48.7	488	2	Q4WRB6	ASPPU	Q4wrB6	aspergillus
116	38	50.0	2223	2	Q5TNJ2	ANOAG	Q5tnj2	anopheles g	189	37	48.7	490	2	Q7RXX3	NEUCR	Q7rxX3	neutropori
117	38	50.0	2493	2	Q6CMV4	KLULA	Q6cmv4	kluyveromyc	190	37	48.7	491	2	Q9HN17	HALSA	Q9hn17	halobacteri
118	38	50.0	6540	2	Q4Q9Y2	LEIMA	Q4q9y2	leishmania	191	37	48.7	492	1	YGFH	ECOLI	YGFH	escherichia
119	37.5	49.3	293	2	Q651G4	ORYSA	Q651g4	oryza sativ	192	37	48.7	492	2	Q8XD13	ECOS7	Q8xd13	coxiella bu
120	37	48.7	74	2	Q9W1W3	DROME	Q9w1w3	drosophila	193	37	48.7	496	2	Q83D15	COXBU	Q83d15	magnaporthe
121	37	48.7	82	2	Q4V1H1	BACCP	Q4v1h1	bacillus ce	194	37	48.7	500	2	Q51N53	MAGGR	Q51n53	corynebacte
122	37	48.7	128	2	Q546B0	STRPY	Q546b0	streptococc	195	37	48.7	502	2	Q8FS81	COREF	Q8fs81	brucella su
123	37	48.7	141	2	Q7X0Q2	STRPY	Q7x0q2	streptococc	196	37	48.7	502	2	Q8FV55	BRUSU	Q8fv55	geobacillus
124	37	48.7	141	2	Q7X0R1	STRPY	Q7x0r1	streptococc	197	37	48.7	503	1	CORO	GEOKA	Q7w7u8	bordetella
125	37	48.7	141	2	Q7X0S8	STRPY	Q7x0s8	streptococc	198	37	48.7	505	2	Q7W7U8	BORPA	Q7w7u8	bordetella
126	37	48.7	141	2	Q7X0T2	STRPY	Q7x0t2	streptococc	199	37	48.7	505	2	Q7WL88	BORBR	Q7wl88	brucella ab
127	37	48.7	141	2	Q7X0T8	STRPY	Q7x0t8	streptococc	200	37	48.7	510	2	Q579X8	BRUAB	Q579x8	brucella ab
128	37	48.7	141	2	Q7X0V3	STRPY	Q7x0v3	streptococc	201	37	48.7	510	2	Q7VMZ7	BORPE	Q7vmz7	clostridium
129	37	48.7	154	2	Q51BB3	ENTHI	Q51bb3	entamoeba h	202	37	48.7	510	2	CAT1	CLQKL	Q36616	echinosacch
130	37	48.7	185	2	Q9F5P6	BRUAB	Q9f5p6	brucella ab	203	37	48.7	544	1	DSK1	SCHPO	Q821l0	salmonella
131	37	48.7	185	2	Q4HRK2	CAMUP	Q4hrk2	campylobact	204	37	48.7	571	2	Q8ZLH0	SALTY	Q8zlh0	brucella me
132	37	48.7	188	2	Q9T0M0	BRAOL	Q9t0m0	brassica ol	205	37	48.7	584	2	Q8YD77	BRUME	Q8ydf7	cryptococcu
133	37	48.7	189	2	Q4QCR9	LEIMA	Q4qcr9	leishmania	206	37	48.7	619	2	Q5KFN8	CRYNE	Q5kfn8	cryptococcu
134	37	48.7	202	2	Q04712	ARATH	Q04712	arabidopsis	207	37	48.7	619	2	Q55QH8	CRYNE	Q55qh8	eurytelma c
135	37	48.7	209	2	Q894C8	CLOTE	Q894c8	clostridium	208	37	48.7	628	1	HCYG	EJRCA	Q9nf14	eurytelma c
136	37	48.7	212	2	Q4V7K1	XENLA	Q4v7k1	xenopus lae	209	37	48.7	628	1	Q6R5P6	MOUSE	Q6r5p6	mus musculus
137	37	48.7	217	2	Q844Q1	9BACT	Q844q1	uncultured	210	37	48.7	652	2	Q9Z117	MOUSE	Q9z117	mus musculus
138	37	48.7	225	2	Q5C6N3	SCHJA	Q5c6n3	schistosoma	211	37	48.7	667	2	Q4HGZ5	CIOIN	Q4hgZ5	ciona ince
139	37	48.7	225	2	Q3ZSP5	LEGPH	Q3zsf5	legionella	212	37	48.7	680	2	Q4UGK6	THEAN	Q4ugk6	theileria a
140	37	48.7	229	2	Q4FR66	9GAMM	Q4fre6	psychobact	213	37	48.7	686	2	Q51S28	MAGGR	Q51s28	magnaporthe
141	37	48.7	232	2	Q9CUB6	MOUSE	Q9cub6	mus musculus	214	37	48.7	709	2	Q6SA09	MOUSE	Q6sa09	mus musculus
142	37	48.7	256	2	Q4NYN1	9DELTA	Q4nyn1	anaeromyxob	215	37	48.7	719	2	Q756M5	ASHGO	Q756m5	ashbva goss
143	37	48.7	261	2	Q871V0	VIBPA	Q871v0	vibrio para	216	37	48.7	732	2	Q9SKB1	ARATH	Q9skb1	arabidopsis
144	37	48.7	269	2	Q81U20	HUMAN	Q81u20	homo sapien	217	37	48.7	737	1	CATA	YERPE	Q9x6b0	yersinia pe
145	37	48.7	272	1	YFAP1	YBAST	Q12331	saccharomyc	218	37	48.7	758	2	Q66E82	YERPS	Q66e82	yersinia ps
146	37	48.7	272	2	Q1LFP1	9BURK	Q1lpf4	burkholderi	219	37	48.7	758	2	Q16610	CAEEL	Q16610	caenorhabdi
147	37	48.7	273	2	Q812B0	RALSO	Q8y2b0	raistonia s	220	37	48.7	808	1	DRP3A	ARATH	DRP3A	haemophilus
148	37	48.7	295	2	Q4X2X7	PLACH	Q4x2x7	plasmodium	221	37	48.7	821	1	PHSG	HAEIN	Q9hd44	trichoderma
149	37	48.7	301	2	Q21274	CAEEL	Q21274	caenorhabdi	222	37	48.7	838	2	Q9HDT4	TRIIE	Q9hdt4	trichoderma
150	37	48.7	301	2	Q6ESX9	ORYSA	Q6esx9	oryza sativ	223	37	48.7	847	2	Q043B3	BRAOL	Q043b3	brassica ol
151	37	48.7	316	2	Q6D6T7	BRWCT	Q6d6t7	erwinia car	224	37	48.7	847	2	P93068	BRAOL	P93068	brassica ol
152	37	48.7	316	2	Q74LX2	LACJO	Q74lx2	lactobacill	225	37	48.7	922	2	Q80217	MOUSE	Q80217	mus musculus
153	37	48.7	341	2	Q6EEC1	PRODO	Q6eec1	protopterus	226	37	48.7	938	2	Q8CIA2	MOUSE	Q8cia2	mus musculus
154	37	48.7	345	2	Q9HBL6	HUMAN	Q9hbl6	homo sapien	227	37	48.7	1043	2	Q5AXS3	EMENI	Q5axs3	aspergillus
155	37	48.7	350	2	Q73CW3	BACCL	Q73cw3	pseudomonas	228	37	48.7	1063	2	Q6LFG6	PLAF7	Q6lfg6	plasmodium
156	37	48.7	352	1	AVRC	PSESG	P13836	pseudomonas	229	37	48.7	1177	2	Q7RFX8	PLAY7	Q7rfx8	plasmodium
157	37	48.7	352	2	Q52378	PSEEX	Q52378	pseudomonas	230	37	48.7	1185	2	Q57114	MOUSE	Q57114	mus musculus
158	37	48.7	353	2	Q4MW16	BACCE	Q4mw16	bacillus ce	231	37	48.7	1274	2	Q57114	MOUSE	Q57114	mus musculus
159	37	48.7	363	2	Q5E857	VIBF1	Q5e857	vibrio fisc	232	37	48.7	1275	2	Q61K10	CAEBR	Q61k10	caenorhabdi
160	37	48.7	371	2	Q5QG14	PSESH	Q5qg14	pseudomonas	233	37	48.7	1455	2	Q61K10	CAEBR	Q61k10	caenorhabdi
161	37	48.7	375	2	Q9JWD3	NEIMA	Q9jwd3	neisseria m	234	37	48.7	1710	2	Q9V9N7	DROME	Q9v9n7	drosophila
162	37	48.7	383	2	Q6CY54	KLULA	Q6cy54	kluyveromyc	235	37	48.7	2008	2	Q4KWH4	BARHE	Q4kwh4	bartonella
163	37	48.7	384	2	Q4LNVS	9BURK	Q4lnv5	burkholderi	236	37	48.7	2082	2	Q6G2D3	BARHE	Q6g2d3	bartonella
164	37	48.7	385	2	Q65YF2	HUMAN	Q65yf2	homo sapien	237	37	48.7	2017	2	Q4QDI3	LEIMA	Q4qdi3	leishmania
165	37	48.7	388	2	Q45146	CAEEL	Q45146	caenorhabdi	238	37	48.7	2060	2	Q9P980	PICAN	Q9p980	pichia angu
166	37	48.7	391	2	Q95K51	MACFA	Q95k51	macaca fasc	239	37	48.7	2414	2	Q6FQZ4	CANGA	Q6fqz4	candida gla
167	37	48.7	395	2	Q4ZRR0	PSESY	Q4zrr0	pseudomonas	240	37	48.7	2479	2	Q7SDD4	NEUCR	Q7sdd4	neutropori
168	37	48.7	395	2	Q87X22	PSESM	Q87x22	pseudomonas	241	37	48.7	2556	2	Q6MVS1	NEUCR	Q6mvs1	neutropori
169	37	48.7	398	2	Q51BQ2	BACFN	Q51bq2	bacteroides	242	37	48.7	2860	2	Q4N8K5	THEPA	Q4n8k5	theileria p
170	37	48.7	398	2	Q64SM2	BACFR	Q64sm2	bacteroides	243	37	48.7	3331	2	Q52F95	MAGGR	Q52f95	magnaporthe
171	37	48.7	408	2	Q81488	PLAF7	Q81488	plasmodium	244	37	48.7	3412	2	Q87PW3	VIBPA	Q87pw3	vibrio para
172	37	48.7	416	1	SSR3	HUMAN	Q6bpg4	debaromyce	245	37	48.7	4498	2	Q9W2Z3	DROME	Q9w2z3	drosophila
173	37	48.7	418	2	Q53ZR7	HUMAN	P12745	homo sapien	246	37	48.7	4578	2	Q4UAE4	THEAN	Q4uae4	theileria a
174	37	48.7	427	2	Q9Y070	PERAM	Q9y070	periplaneta	247	36.5	48.0	591	2	Q86T31	HUMAN	Q86t31	homo sapien
175	37	48.7	427	2	Q55V10	CRINE	Q55v10	cryptococcu	248	36.5	48.0	628	2	Q6ZN36	HUMAN	Q6zn36	homo sapien
176	37	48.7	428	2	Q5KL72	CRINE	Q5kl72	cryptococcu	249	36.5	48.0	637	2	Q8NDQ6	HUMAN	Q8ndq6	homo sapien
177	37	48.7	428	2					250	36.5	48.0	660	2				

251	36.5	48.0	678	2	Q94BN9	ARATH	Q94bn9	arabidopsis	324	36	47.4	436	2	Q9M608	CITUP	Q9m608	citrus unsh
252	36.5	48.0	678	2	Q9LM25	ARATH	Q9lm25	arabidopsis	325	36	47.4	436	2	Q9XG7	CITUP	Q9xg7	citrus para
253	36.5	48.0	1434	2	Q7QHC2	ANOGA	Q7qhc2	anopheles g	326	36	47.4	436	2	Q85660	RHIZ	Q85660	bartonella
254	36	47.4	105	2	Q68893	STRPY	Q68893	streptococc	327	36	47.4	439	2	Q7NUA4	CHRW	Q7nuA4	chromobacte
255	36	47.4	119	2	Q8CD64	MOUSE	Q8cd64	mus musculus	328	36	47.4	441	2	Q96DL8	HUMAN	Q96dl8	homo sapien
256	36	47.4	122	2	Q7QCN0	STRPY	Q7qcn0	streptococc	329	36	47.4	442	2	Q6PK81	HUMAN	Q6pk81	homo sapien
257	36	47.4	133	2	Q5L6E6	CHLAB	Q5l6e6	chlamydophi	330	36	47.4	449	2	Q5GH68	MOUSE	Q5gh68	mus musculus
258	36	47.4	134	2	Q7RZ26	NEUCR	Q7rzz26	neurospora	331	36	47.4	450	2	Q5XJA2	BRARE	Q5xja2	brachydanio
259	36	47.4	144	2	Q4RDI8	TETNG	Q4rdi8	tetradodon n	332	36	47.4	456	2	Q9N385	DEBHA	Q9n385	caenorhabdi
260	36	47.4	147	2	Q51HJ3	MAGRY	Q51hj3	magnaporthe	333	36	47.4	457	2	Q6BIE8	DEBHA	Q6bie8	debaromyce
261	36	47.4	159	2	Q4VLT4	PLABE	Q4vlt4	plasmodium	334	36	47.4	467	1	UNC86	CABEL	P13528	caenorhabdi
262	36	47.4	173	2	Q98NK2	RHIL0	Q98nk2	rhizobium 1	335	36	47.4	468	2	Q8JKN0	9VIRU	Q8jkn0	heliothis z
263	36	47.4	177	2	Q58M30	9CAUD	Q58m30	cyanophaga	336	36	47.4	469	2	Q6DCB9	XENLA	Q6dcb9	xenopus lae
264	36	47.4	177	2	Q8D4L8	VIBVU	Q8d4l8	vibrio vuln	337	36	47.4	472	2	Q33N21	9RHIZ	Q33n21	homo sapien
265	36	47.4	185	2	Q4HKC0	CAMLA	Q4hkc0	campylobact	338	36	47.4	483	2	Q5S210	HUMAN	Q5s210	homo sapien
266	36	47.4	185	2	Q5P2F0	AZOSE	Q5p2f0	azocarcus sp	339	36	47.4	492	2	Q97286	PLAF7	Q97286	plasmodium
267	36	47.4	194	2	Q8K3S6	MOUSE	Q8k3s6	mus musculus	340	36	47.4	493	2	Q4UBT2	THEAN	Q4ubt2	theileria a
268	36	47.4	196	2	P72891	SINY3	P72891	synecocyst	341	36	47.4	495	2	Q7KWV2	DICDI	Q7kwv2	dictyostelli
269	36	47.4	203	2	Q84R02	ORYSA	Q84r02	oryza sativ	342	36	47.4	501	2	Q4JXC0	CORJK	Q4jxc0	corynebacte
270	36	47.4	203	2	Q94LH6	ORYSA	Q94lh6	oryza sativ	343	36	47.4	502	2	Q6NUJ5	HUMAN	Q6nuj5	homo sapien
271	36	47.4	207	2	Q7MG59	VIBVU	Q7mg59	vibrio vuln	344	36	47.4	524	2	Q4HUS6	GIBZE	Q4hus6	gibbereila
272	36	47.4	208	2	Q684A3	9VIRU	Q684a3	sulfolobus	345	36	47.4	524	2	Q4FZM6	XENLA	Q4fzm6	xenopus lae
273	36	47.4	220	1	ATPO	NEUCR	Q9p602	neurospora	346	36	47.4	532	2	Q4YVY8	PLABE	Q4yvy8	plasmodium
274	36	47.4	244	2	Q87EE8	XYLFT	Q87ee8	xyliella fas	347	36	47.4	546	2	Q5RA50	PONPY	Q5ra50	pongo pygma
275	36	47.4	244	2	Q87CD8	XYLFT	Q87cd8	xyliella fas	348	36	47.4	551	2	Q4RM16	TETNG	Q4rm16	tetradodon n
276	36	47.4	269	2	Q9KP13	VIBCH	Q9kp13	vibrio chol	349	36	47.4	552	1	ILVD	CLOAB	Q97ee3	clostridium
277	36	47.4	279	2	Q62QX5	HUMAN	Q62qx5	homo sapien	350	36	47.4	552	1	Q4SIA2	TETNG	Q4sia2	tetradodon n
278	36	47.4	280	2	Q4FN98	GRICK	Q4fn98	candidateus	351	36	47.4	562	2	ILVD	SULTO	Q97yko	sulfolobus
279	36	47.4	281	2	Q4XFP9	PLACH	Q4xfp9	plasmodium	352	36	47.4	562	2	Q571K0	MOUSE	Q571k0	mus musculus
280	36	47.4	282	2	Q97LE9	CLOAB	Q97le9	clostridium	353	36	47.4	578	2	Q9RZS2	DEIRA	Q9rzs2	deinococcu
281	36	47.4	290	2	Q8DE68	VIBVU	Q8de68	vibrio vuln	354	36	47.4	579	2	Q521W0	MAGR	Q521w0	magnaporthe
282	36	47.4	290	2	Q7MPF8	VIBVU	Q7mpf8	vibrio vuln	355	36	47.4	579	2	Q87IM8	VIBPA	Q87im8	vibrio para
283	36	47.4	295	2	Q5P6H4	AZOSE	Q5p6h4	azocarcus sp	356	36	47.4	588	2	Q7YZG7	CABEL	Q7yzg7	caenorhabdi
284	36	47.4	295	2	Q8DQ63	STRP6	Q8dq63	streptococc	357	36	47.4	595	2	Q85872	SPHAR	Q85872	spingomona
285	36	47.4	295	2	Q8P9C6	STRPN	Q8p9c6	streptococc	358	36	47.4	596	2	Q01880	CABEL	Q01880	caenorhabdi
286	36	47.4	297	2	Q7X4J9	STRSU	Q7x4j9	streptococc	359	36	47.4	597	2	Q89B14	BRABA	Q89b14	bradyrhizob
287	36	47.4	301	2	Q5V7F5	HALMA	Q5v7f5	haloarcula	360	36	47.4	600	2	Q8GE11	HELMO	Q8ge11	heliobacill
288	36	47.4	304	2	Q51LOU9	GEOKA	Q51lou9	geobacillus	361	36	47.4	603	2	Q8PD49	USTMA	Q8pd49	ustilago na
289	36	47.4	305	2	Q97H99	CLOAB	Q97h99	clostridium	362	36	47.4	603	2	Q559K3	DICDI	Q559k3	dictyostelli
290	36	47.4	308	2	Q568S1	BRARE	Q568s1	brachydanio	363	36	47.4	607	2	Q96M81	HUMAN	Q96m81	homo sapien
291	36	47.4	310	1	PYRB	ECOL6	P0a788	escherichia	364	36	47.4	618	2	Q4R322	MACFA	Q4r322	macaca fasc
292	36	47.4	310	1	PYRB	ECOL6	P0a787	escherichia	365	36	47.4	620	1	YGL8	YEAST	P53207	saccharomyc
293	36	47.4	310	1	PYRB	SHIFL	P0a786	escherichia	366	36	47.4	631	2	Q6GRU3	KULUA	Q6gru3	kluveromyc
294	36	47.4	310	1	Q7NP12	GLOVI	P0a789	shigella fl	367	36	47.4	631	2	Q46020	CABEL	Q46020	caenorhabdi
295	36	47.4	315	2	Q4ZLY4	PSESY	Q4zly4	pseudomonas	368	36	47.4	637	2	Q5AC11	CANAL	Q5ac11	candida alb
296	36	47.4	316	2	Q5GT48	WOLTR	Q5gt48	woibachia s	369	36	47.4	661	2	Q4N948	THEPA	Q4n948	theileria p
297	36	47.4	318	2	Q4G3B7	EMTHU	Q4g3b7	emiliania h	370	36	47.4	663	2	Q617K0	CAEBR	Q617k0	caenorhabdi
298	36	47.4	318	2	Q56K15	BOVIN	Q56k15	bos taurus	371	36	47.4	686	2	Q5U190	DROME	Q5u190	drosophila
299	36	47.4	319	2	Q5UY62	HALMA	Q5uy62	haloarcula	372	36	47.4	691	2	Q54098	STRCO	Q54098	streptomyce
300	36	47.4	320	2	Q5SPP0	BRARE	Q5spp0	brachydanio	373	36	47.4	725	2	Q6G094	BAROU	Q6g094	bartonella
301	36	47.4	329	2	Q86LZ7	BRALA	Q86lz7	branchiosto	375	36	47.4	730	2	Q8T867	DICDI	Q8t867	dictyostelli
302	36	47.4	340	2	Q8JJK7	9VIRU	Q8jjk7	heliothis z	376	36	47.4	733	1	YR2	YEAST	P43602	saccharomyc
303	36	47.4	340	2	Q5CJG3	CRYHO	Q5cjg3	cryptospori	377	36	47.4	758	2	Q5UPF7	HUMAN	Q5upf7	homo sapien
304	36	47.4	351	2	Q5CJG3	CRYHO	Q5cjg3	cryptospori	378	36	47.4	761	2	Q6FW88	CANGA	Q6fw88	candida gla
305	36	47.4	352	2	Q5CS90	CRYPV	Q5cs90	cryptospori	379	36	47.4	763	2	Q7PR04	ANOGA	Q7pr04	anopheles g
306	36	47.4	353	2	Q81HD6	BACCR	Q81hd6	bacillus ce	380	36	47.4	779	2	Q5B718	EMENI	Q5b718	aspergillus
307	36	47.4	360	2	Q87C12	XYLFT	Q87c12	xyliella fas	381	36	47.4	794	2	Q5S8F4	CANAL	Q5s8f4	candida alb
308	36	47.4	368	2	Q5V1Y7	HALMA	Q5v1y7	haloarcula	382	36	47.4	821	2	Q5S8F4	CANAL	Q5s8f4	homo sapien
309	36	47.4	368	2	Q75F62	ASHGO	Q75f62	ashbya gos	383	36	47.4	835	2	Q5G6E4	HUMAN	Q5g6e4	homo sapien
310	36	47.4	369	2	Q6CY5	KLUUA	Q6cy5	kluveromyc	384	36	47.4	875	2	Q5B1L8	EMENI	Q5b1l8	aspergillus
311	36	47.4	369	2	Q72T85	LEPIC	Q72t85	leptospora	385	36	47.4	905	1	DLG1	MOUSE	Q811d0	mus musculus
312	36	47.4	369	2	Q8F256	LEPIN	Q8f256	leptospora	386	36	47.4	915	2	Q6MT28	MYCMS	Q6mt28	mycoplasma
313	36	47.4	374	2	Q4F1V1	USTMA	Q4f1v1	ustilago na	387	36	47.4	925	2	Q75B60	ASHGO	Q75b60	ashbya gos
314	36	47.4	376	2	Q637U6	BACCC	Q637u6	bacillus ce	388	36	47.4	964	2	Q4N3B5	THEPA	Q4n3b5	theileria p
315	36	47.4	379	2	Q8AAV6	BACTN	Q8aav6	bacteroides	389	36	47.4	989	2	Q4N3B5	THEPA	Q4n3b5	theileria p
316	36	47.4	390	2	Q9ZGP9	BUTFI	Q9zgp9	butyrivibri	390	36	47.4	989	2	Q9LTL5	ARATH	Q9ltl5	arabidopsis
317	36	47.4	392	2	Q9S2Y5	STRCO	Q9s2y5	streptomyce	391	36	47.4	989	2	Q9M1E4	ARATH	Q9m1e4	arabidopsis
318	36	47.4	396	2	Q6VLU2	9RHIZ	Q6vlu2	bartonella	392	36	47.4	1014	2	Q75F59	ASHGO	Q75f59	ashbya gos
319	36	47.4	398	2	Q82A75	STRAM	Q82a75	streptomyce	393	36	47.4	1042	2	Q6Q181	RAT	Q6q181	rattus norv
320	36	47.4	399	2	Q41ER5	GIBZE	Q41er5	gibbereila	394	36	47.4	1044	2	Q4WHK0	ASPFU	Q4whk0	aspergillus
321	36	47.4	417	2	Q6CWN3	KJULA	Q6cwn3	kluveromyc	395	36	47.4	1065	2	Q7SEJ3	NEUCR	Q7sej3	neurospora
322	36	47.4	422	2	Q6C0R4	YARLI	Q6c0r4	yarrowia li	396	36	47.4	1066	1	ST55	SCHPO	Q7454	schizosacch
323	36	47.4	423	2	Q96F43	HUMAN	Q96f43	homo sapien	396	36	47.4	1066	1	ST55	SCHPO	Q7454	schizosacch

397	36	47.4	1233	2	Q80w41_MOUSE
398	36	47.4	1230	2	Q4WJ79_ASPFU
399	36	47.4	1250	1	SSD1_YEAST
400	36	47.4	1235	2	Q4F3Z1_USTMA
401	36	47.4	1266	2	Q4WLUO_ASPFU
402	36	47.4	1243	2	Q9U2M6_CABEL
403	36	47.4	1345	2	Q7K714_CAEEL
404	36	47.4	1365	2	Q8UZ48_9VIRU
405	36	47.4	1366	2	Q91MQ5_ARATH
406	36	47.4	1417	2	Q6C661_YARLI
407	36	47.4	1491	2	Q4SCU8_TETNG
408	36	47.4	1505	2	Q6ZQK0_MOUSE
409	36	47.4	1572	2	Q4PHA0_USTMA
410	36	47.4	1679	1	YMF9_YEAST
411	36	47.4	2031	2	Q811R4_PLAF7
412	36	47.4	2114	2	Q9VVA4_DROME
413	36	47.4	2653	2	Q7QSJ2_GIALA
414	35.5	46.7	209	2	Q7SE21_NEUCR
415	35.5	46.7	215	2	Q5NQU9_ZYMO
416	35.5	46.7	303	2	Q8TZ54_METKA
417	35.5	46.7	310	2	Q8YZN7_ANASP
418	35.5	46.7	548	2	Q5CS08_CVIRU
419	35.5	46.7	683	2	Q9J0X7_9P1PV
420	35.5	46.7	2032	2	Q81TH9_NEUCR
421	35	46.1	80	2	Q81HN9_PLAF7
422	35	46.1	91	2	Q7O2Y7_GIALA
423	35	46.1	105	2	Q5X4M2_LEGPA
424	35	46.1	105	2	Q7MBA0_PHOLL
425	35	46.1	112	2	Q7NAV4_PHOLL
426	35	46.1	115	2	Q5PH53_EHRRG
427	35	46.1	112	2	Q5HC24_EHRRW
428	35	46.1	113	1	RL31_CANGA
429	35	46.1	113	2	Q6BS72_DEBHA
430	35	46.1	115	2	Q7Y3Y5_9CAUD
431	35	46.1	116	2	Q6C4Q8_YARLI
432	35	46.1	118	2	Q4K7Q8_BACHD
433	35	46.1	121	2	Q4T6D3_TETNG
434	35	46.1	124	2	Q5MIR2_ABDAL
435	35	46.1	132	2	Q8CTZ0_STAEP
436	35	46.1	135	1	RL32_NANQO
437	35	46.1	147	2	Q4LY59_9BURK
438	35	46.1	150	2	Q9ABE0_CAUDR
439	35	46.1	153	2	Q4NLJ3_9MICC
440	35	46.1	158	2	Q7T3H0_BRARE
441	35	46.1	160	2	Q5MVS1_LEGPA
442	35	46.1	160	2	Q5X4D6_LEGPA
443	35	46.1	161	2	Q5ZUL7_LEGPH
444	35	46.1	164	1	MR4Z_GJUXO
445	35	46.1	167	2	Q7Q899_ANOGA
446	35	46.1	168	1	VE6_HP721
447	35	46.1	172	1	ADXH_DROME
448	35	46.1	172	1	VP19_HARV
449	35	46.1	172	1	VP19_PLV
450	35	46.1	172	2	Q4VM12_9TOMB
451	35	46.1	173	1	VP19_CNV
452	35	46.1	173	2	Q91BS5_CNV
453	35	46.1	173	2	Q91BS6_CNV
454	35	46.1	173	2	Q91BS7_CNV
455	35	46.1	173	2	Q91ES8_CNV
456	35	46.1	173	2	Q91ES9_CNV
457	35	46.1	174	2	Q9CRZ6_MOUSE
458	35	46.1	180	1	DCD_AQJAE
459	35	46.1	183	2	Q8KEB3_RHLET
460	35	46.1	188	2	Q7RYU0_PLAYO
461	35	46.1	192	2	Q6H5X0_ORYSA
462	35	46.1	199	1	Y396_BUCAP
463	35	46.1	208	2	Q651Z1_RANFI
464	35	46.1	202	2	Q7Z202_BRARE
465	35	46.1	215	2	Q6G108_XENTR
466	35	46.1	225	2	Q5QXM0_IDILO
467	35	46.1	229	2	Q5YVC2_NOCFA
468	35	46.1	230	2	Q97Z12_SULSO
469	35	46.1	232	2	Q6G

470	35	46.1	232	2	Q6GGC8	STAAR	staphylococ
471	35	46.1	232	2	QSHF18	STAAR	staphylococ
472	35	46.1	232	2	Q7A0S2	STAAC	staphylococ
473	35	46.1	232	2	Q7A5C6	STAAN	staphylococ
474	35	46.1	232	2	Q97TS4	STAAM	staphylococ
475	35	46.1	246	2	Q86DY0	SCHJA	schistosoma
476	35	46.1	266	2	Q60AD9	METCA	methyllococ
477	35	46.1	266	2	Q5V7K8	HALMA	haloarcula
478	35	46.1	270	2	Q96MC9	HUMAN	homo sapien
479	35	46.1	270	2	Q5JTM5	HUMAN	homo sapien
480	35	46.1	271	2	Q9KG55	BACHD	bacillus ha
481	35	46.1	280	2	Q5E835	V18F1	Q5E835 bacillus fiac
482	35	46.1	280	2	Q6SH12	9BACT	Q6SH12 uncultured
483	35	46.1	281	1	MURB_FUSN		Q8rdq3 fusobacteri
484	35	46.1	281	2	Q54XH6	D1CDI	Q54XH6 dictyostelli
485	35	46.1	283	2	Q8E051	STRA5	Q8E051 streptococc
486	35	46.1	283	2	Q8E558	STRA3	Q8E558 streptococc
487	35	46.1	287	2	Q6WS00	ENTFA	Q6ws00 enterococc
488	35	46.1	292	1	KHSE_FYRUT		Q8u1z9 pyrococcus
489	35	46.1	295	2	Q9CPC9	PASMU	Q9cpc9 pasteurella
490	35	46.1	297	2	Q6CR68	KUOLA	Q6cr68 kluyveromyc
491	35	46.1	297	2	Q51755	BORBU	Q51755 borrelia bu
492	35	46.1	301	2	O27842	MYTHX	O27842 methanobact
493	35	46.1	301	2	Q846R0	MYXHA	Q846r0 myxococcu
494	35	46.1	306	2	Q6BZC6	DBBHA	Q6bzc6 debaryomyce
495	35	46.1	308	2	Q7BSB3	ASHGO	Q7bsb3 ashbya goss
496	35	46.1	309	2	Q8R9R1	THETN	Q8-r9r1 thermoanaer
497	35	46.1	312	1	FMT_MYCTO		F64135 mycobacteri
498	35	46.1	312	1	FMT_MYCTO		P64134 mycobacteri
499	35	46.1	314	1	MR2_YEAST		P21651 saccharomyc
500	35	46.1	315	2	Q6FNV3	CANGA	Q6fnv3 candida gla

ALIGNMENTS

RESULT 1

SEPR_HUMAN STANDARD; PRT: 760 AA.
 ID Q12884; Q00199; Q86229; Q99998; Q9UD4;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Sepsiae (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
 DE membrane serine protease) (170-kDa melanoma membrane-bound
 DE gelatinase).
 GN Name=FAP;
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 NCBI_TaxID=9606;
 [1]
 RN NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RP TISSUE=Fibroblast;
 RX MEDLINE=94261645; PubMed=7911242;
 RA Scanlan M.J., Raj B.K.M., Calvo B., Garin-Chesa P., Sanz-Moncasi M.P.,
 RA Healey J.H., Old L.J., Rettig W.J.;
 RT "Molecular cloning of fibroblast activation protein alpha, a member of
 RT the serine protease family selectively expressed in stromal
 RT fibroblasts of epithelial cancers."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5657-5661(1994).
 [2]
 RN NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT THR-354.
 RP TISSUE=Melanoma;
 RX MEDLINE=97388251; PubMed=9247085; DOI=10.1016/S0925-4439(97)00032-X;
 RA Goldstein D.A., Gherif G., Pineiro-Sanchez M.L., Salamone M., Yeh Y.,
 RA Fleissate L., Chen W.-T.;
 RT "Molecular cloning of seprase: a serine integral membrane protease
 RT from human melanoma."
 RL Biochim. Biophys. Acta 1361:11-19(1997).
 [3]
 RN NUCLEOTIDE SEQUENCE (ISOFORM 1), AND PROTEIN SEQUENCE OF 220-229;
 RP

461-472 AND 511-518.
 CC TISSUE=Melanoma;
 RX MEDLINE=97218181; PubMed=9065413; DOI=10.1074/jbc.272.12.7595;
 RA Pineiro-Sanchez M.L., Goldstein L.A., Dodd J., Howard L., Yeh Y.,
 Chen W.-T.;
 RT Identification of the 170-kDa melanoma membrane-bound gelatinase
 (seprase) as a serine integral membrane protease.";
 RL J. Biol. Chem. 272:7595-7601(1997).
 [4]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
 CC TISSUE=Melanoma;
 RX MEDLINE=20112818; PubMed=10644713; DOI=10.1074/jbc.275.4.2554;
 RA Goldstein L.A., Chen W.-T.;
 RT "Identification of an alternatively spliced seprase mRNA that encodes
 a novel intracellular isoform.";
 RL J. Biol. Chem. 275:2554-2559(2000).
 [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
 CC THR-354.
 RX TISSUE=Placenta;
 CC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
 Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullaly S.J.,
 Bask S.A., McEwan P.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [6]
 RP PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.
 RX MEDLINE=94327249; PubMed=7519584;
 RA Rettig W.J., Su S.L., Fortunato S.R., Scanlan M.J., Raj B.K.M.,
 Garin-Chesa P., Healey J.H., Old L.J.;
 RT "Fibroblast activation protein: purification, epitope mapping and
 induction by growth factors.";
 RL Int. J. Cancer 58:385-392(1994).
 CC -!- FUNCTION: May have a role in tissue remodeling during development
 and wound healing, and may contribute to invasiveness in malignant
 cancers.
 CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
 type IV collagen, but not native type I or type IV collagen. Does
 not cleave laminin, fibronectin, fibrin or casein.
 CC -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
 inactive.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell
 surface lamellipodia, invadopodia and on shed vesicles.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=L;
 CC IsoId=Q12884-1; Sequence=Displayed;
 CC Note=Major isoform;
 CC Name=2; Synonyms=S, Truncated;
 CC IsoId=Q12884-2; Sequence=VSP_005367;
 CC -!- TISSUE SPECIFICITY: Fibroblast-specific.
 CC -!- INDUCTION: In fibroblasts at times and sites of tissue remodeling
 during development, tissue repair, and carcinogenesis.
 CC -!- PTM: N-glycosylated.
 CC -!- PTM: The N-terminus may be blocked.
 CC -!- SIMILARITY: Belongs to the peptidase S9B family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

EMBL; U09278; AAB49652.1; -; mRNA.
 DR EMBL; U76833; AAC51668.1; -; mRNA.
 DR EMBL; AF007822; AAF21600.1; -; mRNA.
 DR EMBL; BC026250; AAH26250.1; -; mRNA.
 DR PDB; 1Z68; X-ray; A=39-757, B=-.
 DR MEROPS; S09.007; -; -.
 DR Ensembl; ENSG00000078098; Homo sapiens.
 DR HGNC; HGNC:3590; FAP.
 DR MIM; 600403; -.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0030027; C:lamellipodium; IDA.
 DR GO; GO:0005886; C:plasma membrane; NAS.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.
 DR GO; GO:0004222; F:metalloendopeptidase activity; TAS.
 DR GO; GO:0042803; F:protein homodimerization activity; NAS.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_sptrs.
 DR Pfam; PF00930; DPPIV N; 1.
 DR Pfam; PF00326; Peptidase S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW 3D-structure; Alternative splicing; Direct protein sequencing;
 KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
 KW Signal-anchor; Transmembrane.
 FT TOPO_DOM 1 4
 FT TRANSMEM 5 25
 FT Cytoplasmic (Potential).
 FT Signal-anchor for type II membrane
 FT protein (Potential).
 FT Extracellular (Potential).
 FT Charge relay system (by similarity).
 FT Charge relay system (by similarity).
 FT Charge relay system (by similarity).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 FT Missing (in isoform 2).
 FT /FTid=VSP_005367.
 FT R -> T (in dbSNP:1126507).
 FT /FTid=VAR_019691.
 FT P -> A (in Ref. 2 and 5).
 FT K -> T (in Ref. 2 and 5).
 SQ SEQUENCE 760 AA; 87821 MW; A0D34B4801BE07EA CRC64;

Query Match 100.0%; Score 76; DB 1; Length 760;
 Best Local Similarity 100.0%; Pred. No. 8.1e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEET 15
 |||||
 Db 24 IVLRPSRVHNSSEET 38

RESULT 2
 Q53TP5 HUMAN
 ID Q53TP5_HUMAN PRELIMINARY; PRT; 760 AA.
 AC Q53TP5;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein FAP.

GN Name=FAP;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cotton M., Maupin R., Hawkins M., Harkins R.;
RT "The sequence of Homo sapiens BAC clone RP11-576116."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007750; AAY24205.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR PROSITE; PS00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 760 AA; 87113 MW; 7FF817B54F75142 CRC64;

Query Match 100.0%; Score 76; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLRPSRVHSEENT 15
DB 24 IVLRPSRVHSEENT 38
|||||
|||||

RESULT 3
Q6P7D6 RAT PRELIMINARY; PRT; 350 AA.
AC Q6P7D6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Pap protein.
GN Name=Pap;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavaant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RG NIH MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061713; AAH61713.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
SQ SEQUENCE 350 AA; 40869 MW; 49FF2B46A9CF0F7F CRC64;

Query Match 69.7%; Score 53; DB 2; Length 350;
Best Local Similarity 73.3%; Pred. No. 0.55;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IVLRPSRVHSEENT 15
DB 24 IVLRPSRVHSEENT 38
|||||
|||||

RESULT 4
Q8R492 RAT PRELIMINARY; PRT; 761 AA.
AC Q8R492;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fibroblast activation protein alpha subunit.
GN Name=Pap;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zuckschwerdt K., Park J.E.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493782; AAM11677.1; -; mRNA.
DR HSP; P27487; lpfq.
DR MEROPS; S09.007; -.
DR Ensembl; ENSRNOG00000005679; Rattus norvegicus.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR PROSITE; PS00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 761 AA; 87843 MW; EB214BB54BB14F2C CRC64;

Query Match 69.7%; Score 53; DB 2; Length 761;
Best Local Similarity 73.3%; Pred. No. 1.4;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 IVLRPSRVHNSSENT 15
Db 24 IVLRPSRVSPKPGNT 38

RESULT 5
SEPR MOUSE STANDARD; PRT; 761 AA.
AC P97321;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Seprase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
DE membrane serine protease).
GN Name=Tap;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Murcidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).
RC STRAIN=BA1B/c; TISSUE=Embryo;
RX MEDLINE=97284459; PubMed=9139873;
RA Niedermeier J., Scanlan M.J., Garin-Chesa P., Daiber C., Fiebig H.H.,
RA Old L.J., Rettig W.J., Schnapp A.;
RT "Mouse fibroblast activation protein: molecular cloning, alternative
RT splicing and expression in the reactive stroma of epithelial
RT cancers.";
RL Int. J. Cancer 71:383-389(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Mammary gland.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.P., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo A.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usslin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.N., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May have a role in tissue remodeling during development
CC and wound healing, and contribute to invasiveness in malignant
CC cancers.
CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
CC type IV collagen, but not native type I or type IV collagen. Does
CC not cleave laminin, fibronectin, fibrin or casein.
CC -!- SUBUNIT: Homodimer, or heterodimer with DPPA. The monomer is
CC inactive (by similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC isoId=P97321-1; Sequence=Displayed;
CC Name=2;
CC isoId=P97321-2; Sequence=VSP_005368;
CC Name=3;
CC isoId=P97321-3; Sequence=VSP_005369;
CC -!- TISSUE SPECIFICITY: Detected in fibroblasts, in placenta, uterus,
CC embryos from day 7-19 and in new-born mice (P1).

-!- SIMILARITY: Belongs to the peptidase S9B family.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
EMBL; Y10007; CAA71116.1; -, mRNA.
EMBL; BC019190; AAH19190.1; -, mRNA.
HSSP; P27487; 1PFO.
MEROPS; S09.007.
ENEMENI; ENSMUSG00000000392; Mus musculus.
GO; GO:0016021; C: integral to membrane; TAS.
MGI; MGI:109608; Fap.
InterPro; IPR002471; Pept_S9_AS.
InterPro; IPR001375; Peptidase_S9.
InterPro; IPR002469; Peptidase_S9B.
InterPro; IPR000379; Ser_estrse.
Pfam; PF00930; DPIP_N; 1.
PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
Alternative splicing; Glycoprotein; Hydrolase; Protease;
Serine protease; Signal-anchor; Transmembrane.
FT TOPO_DOM 1 4 Cytoplasmic (Potential).
FT TRANSMEM 5 25 Signal-anchor for type II membrane
FT protein (Potential).
FT TOPO_DOM 26 761 Extracellular (Potential).
FT ACT_SITE 624 624 Charge relay system (By similarity).
FT ACT_SITE 702 702 Charge relay system (By similarity).
FT ACT_SITE 734 734 Charge relay system (By similarity).
FT CARBOHYD 49 49 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 92 92 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 227 227 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
FT VARSPPLIC 31 63 Missing (in isoform 3).
FT VARSPPLIC 31 35 Missing (in isoform 2).
FT CONFLICT 737 737 S -> L (in Ref. 2).
SQ SEQUENCE 761 AA; 87945 MW; 9174C3AEDA213B25 CRC64;
Query Match 68.4%; Score 52; DB 1; Length 761;
Best Local Similarity 73.3%; Pred. No. 2.1;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 IVLRPSRVHNSSENT 15
Db 24 IVLRPSRVSPKPGNT 38

RESULT 6
Q73QP6 TREDE PRELIMINARY; PRT; 325 AA.
AC Q73QP6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Oligopeptide/dipeptide ABC transporter, ATP-binding protein.
GN OrderedLocustNames=TDE0397;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,

```

RA Gebregorgis E., Geer K., Tsegaye G., Malek J.A., Ayodeji B.,
RA Shatsman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstock G.M., Norris S.J., Frazer C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen *Treponema denticola*
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
CC -J- SIMILARITY: Belongs to the ABC transporter family.
CC EMBL; AE017247; AAS10892.1; -, Genomic_DNA.
DR TIGR; TDE0397; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0000189; F:oligopeptide transporter activity; IEA.
DR GO; GO:0006857; P:oligopeptide transport; IEA.
DR InterPro; IPR003439; ABC transp. like.
DR InterPro; IPR010066; Oligo HPV_C.
DR Pfam; PF00005; ABC trans. 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR01727; oligo HPV; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome; Membrane; Nucleotide-binding;
KW Peptide transport; Protein transport; Transport.
SQ SEQUENCE 325 AA; 36788 MW; 04A10601D8FAC7CF CRC64;

Query Match 60.5%; Score 46; DB 2; Length 325;
Best Local Similarity 57.1%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEN 14
|:|:|:|:|:|:|
Db 116 ILLRPLKHNIAEN 129
|:|:|:|:|:|:|

RESULT 7
Q4RR26_TETNG PRELIMINARY; PRT; 277 AA.
AC Q4RR26;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 14 SCAF15003, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00030348001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Kalinka M., Vacherie B.,
RA Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutancieu J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -I- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01015003; CAG09156.1; -, Genomic_DNA.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 31740 MW; F5C479D3D0811CC8 CRC64;

Query Match 59.2%; Score 45; DB 2; Length 277;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSEN 14
|:|:|:|:|:|:|
Db 222 VLEENKVHNSDN 234
|:|:|:|:|:|:|

RESULT 8
Q6IJW5_DROME PRELIMINARY; PRT; 81 AA.
AC Q6IJW5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HPC14087.
GN ORFNames=HDC14087;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the *Drosophila* genome.";
RL Genome Biol. 5:RESEARCH003.1-RESEARCH003.17(2003).
CC -I- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK002601; DAA04107.1; -, Genomic_DNA.
SQ SEQUENCE 81 AA; 9217 MW; D73008A7E41527F2 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 81;
Best Local Similarity 58.3%; Pred. No. 4.4;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEN 15
|:|:|:|:|:|:|
Db 59 RPSRLHNSNENS 70
|:|:|:|:|:|:|

RESULT 9
Q61710_CAEBR PRELIMINARY; PRT; 809 AA.
ID Q61710_CAEBR PRELIMINARY;
AC Q61710;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14957.
GN Names=CBG14957;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RP The C.briggsae Sequencing Consortium;
RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL

CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: CAAC01000068; CAB68975.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 809 AA; 91832 MW; 00EB1442D461F33C CRC64;

Query Match 57.9%; Score 44; DB 2; Length 809;
 Best Local Similarity 66.7%; Pred. No. 68;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 14
 ||||| ||||| :
 Db 205 LRPSRVHNSSE 216

RESULT 10
 Q4WUH2 ASPFU

ID Q4WUH2 ASPFU PRELIMINARY; PRT; 862 AA.

AC Q4WUH2

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=Afu508450;

OS Aspergillus fumigatus Af293.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxID=330879;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Af293;

RA Nieman W., Pain A., Anderson M.J., Wortman J., Kim H. Stanley,

RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,

RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,

RA Farman N., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,

RA Foster M., Fraser A., Garcia J.L., Garcia M.J., Goble A.,

RA Goldman G.H., Goni K., Griffith-Jones S., Guilliam R., Haas B.,

RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,

RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,

RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,

RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,

RA Mouny I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,

RA Penvalva M.A., Perteau M., Price C., Pritchard B.L., Quail M.A.,

RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,

RA Rensing C.M., Rukter S., Salzberg S.L., Sanchez M.,

RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,

RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,

RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,

RA Machida M., Hall N., Barrell B., Denning D.W.;

RT "Genomic sequence of the pathogenic and allergenic filamentous fungus

Aspergillus fumigatus";

RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC preliminary data.

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL: AAH01000003; BA01754.1; -; Genomic_DNA.

KW Hypothetical protein.

SQ SEQUENCE 862 AA; 96297 MW; B050E61F036D7768 CRC64;

Query Match 56.6%; Score 43; DB 2; Length 862;
 Best Local Similarity 64.3%; Pred. No. 11e+02;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSE 15
 ||||| ||||| :
 Db 471 VLRPSRVHNSSE 484

RESULT 11

Q5ATIL_EWENI

ID Q5ATIL_EWENI PRELIMINARY; PRT; 1116 AA.

AC Q5ATIL; 2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AN8399.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ranasingh U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vasiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Genome Sequence of Aspergillus nidulans";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AACD01000153; EAA67021.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 1116 AA; 122163 MW; 18B57D2CF3B47973 CRC64;
 Query Match 56.6%; Score 43; DB 2; Length 1116;
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 LRPSRVHNSSE 13
 :|||:|:|:| :
 Db 1093 VRPSRLHGKE 1103
 RESULT 12
 Q4Q3F8_LEIMA
 ID Q4Q3F8_LEIMA PRELIMINARY; PRT; 257 AA.
 AC Q4Q3F8
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Splicing factor pter1 interacting protein, putative.
 GN ORFNames=LmjF34_0495;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
 RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neill S.,
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CT005271; CAJ07753.1; -; Genomic_DNA.

```
SQ SEQUENCE 257 AA; 30326 MW; 968094356013F406 CRC64;
Query Match 55.3%; Score 42; DB 2; Length 257;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEENT 15
Db 212 RSSRVHNSSESS 223

RESULT 13
Q9NKE8_HUMAN PRELIMINARY; PRT; 425 AA.
AC Q9NKE8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein FLJ20291.
GN Name=FLJ20291;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Oiyashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RA Submitted (FEB-2000) to the ENBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RA KAWAKAMI T., NOGUCHI S., ITOH T., SHIGETA K., SENBA T., MATSUMURA K.,
RA NAKAJIMA Y., MIZUNO T., MORINAGA M., OTA T., SUZUKI Y., OIYASHI M.,
RA NISHI T., SHIBAHARA T., TANAKA T., NAKAMURA Y., ISEGAI T., SUGANO S.;
RA Submitted (FEB-2000) to the ENBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haileh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ugin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McWeeney J.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the ENBL/GenBank/DBJ databases.
DR EMBL; AK000298; BAA01065.1; -; mRNA.
DR EMBL; BC008833; AAH08833.1; -; mRNA.
DR Ensembl; ENSG00000108296; Homo sapiens.
KW Hypothetical protein.
SQ SEQUENCE 425 AA; 49647 MW; F8288E143A3FF764 CRC64;
Query Match 55.3%; Score 42; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 74;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEENT 14
Db 237 RPSKLHNSKVN 307
```

```
RESULT 14
Q6BYR6_DEBHA PRELIMINARY; PRT; 435 AA.
AC Q6BYR6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Debaryomyces hansenii chromosome A of strain CBS767 of Debaryomyces
DE hansenii.
GN OrderedLocusNames=DEHA0A07634g;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=4959;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36239 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durkens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykassen C.,
RA Boistrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellanz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Sennen D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
DR EMBL; CR382133; CAG84609.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001005; Myb_DNA_bd.
DR Pfam; PF00249; Myb_DNA-binding; 1.
DR SMART; SM00717; SANT; 1.
DR PROSITE; PS00090; MYB_3; 1.
KW Complete proteome; Nuclear protein; Repeat.
SQ SEQUENCE 435 AA; 46322 MW; 335C47803877657 CRC64;
Query Match 55.3%; Score 42; DB 2; Length 435;
Best Local Similarity 54.5%; Pred. No. 76;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RPSRVHNSSEENT 15
Db 403 PNKTHNSQKNT 413

RESULT 15
Q8YWI4_ANASP PRELIMINARY; PRT; 444 AA.
AC Q8YWI4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE All1802 protein.
GN OrderedLocusNames=all1802;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
```


RA Kishida Y., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
 RA Nakazaki N., Shingo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium *Anabaena* sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; BA000019; BAB73501.1; -; Genomic_DNA.
 DR PIR; AD2031; AD2031.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008324; C:cation transporter activity; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR InterPro; IPR003445; Cat_transpt.
 DR InterPro; IPR004772; K_transpt.
 DR Pfam; PF02386; TKH; 1.
 DR TIGRPFAMs; TIGR00933; 2a38; 1.
 KW Complete proteome.
 SQ SEQUENCE 444 AA; 48154 MW; 729945E40614D03D CRC64;

Query Match 55.3%; Score 42; DB 2; Length 444;
 Best Local Similarity 63.6%; Pred. NO. 78;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPSRVHNSSEEN 14
 |||:|:|
 DB 430 RPTIHYPEEN 440

RESULT 16
 ID Q59NX5 CANAL PRELIMINARY; PRT; 787 AA.
 AC Q59NX5;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Potential fungal zinc cluster transcription factor.
 GN ORFNames=CaO19.6680;
 OS *Candida albicans* SC5314.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Mitosporic Saccharomycetales; *Candida*.
 OX NCBI_TaxID=237561;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SC5314;
 RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
 RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
 RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
 RA Davis R.W., Scherer S.;
 RT "The diploid genome sequence of *Candida albicans*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SC5314;
 RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
 RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
 RA Jones T., Scherer S., Agabian N.;
 RT "Annotation of the Genome of *Candida albicans*.";
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AACQ01000199; EAK92203.1; -; Genomic_DNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
 KW Transcription regulation; Zinc.
 SQ SEQUENCE 787 AA; 89731 MW; 74515755F68D6C CRC64;

Query Match 55.3%; Score 42; DB 2; Length 787;
 Best Local Similarity 53.8%; Pred. NO. 1.5e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLRPSRVHNSSEEN 14

Db 547 VLNPSQISNNEEDN 559
 |||:|:|

RESULT 17
 Q7X125 ORYSA PRELIMINARY; PRT; 837 AA.
 ID Q7X125;
 AC Q7X125;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative S-receptor kinase.
 GN Name=P0506C07.6; Synonyms=QJ1339_B08.15;
 OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; *Oryza*.
 OX NCBI_TaxID=39947;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone:QJ1339_B08.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AP004384; BAC79932.1; -; Genomic_DNA.
 DR EMBL; AP003753; BAD30190.1; -; Genomic_DNA.
 DR Gramene; Q7X125; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001480; B_lectin.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR00719; Pan_app.
 DR InterPro; IPR008271; Ser_thr_kin_AS.
 DR Pfam; PF01453; B_lectin; 1.
 DR Pfam; PF00024; PAN; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00108; B_lectin; 1.
 DR PROSITE; PS0927; BULB_LECTIN; 1.
 DR PROSITE; PS0948; PAN; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Nucleotide-binding; Receptor;
 KW Serine/threonine-protein kinase; transferase.

SQ SEQUENCE 837 AA; 92106 MW; 29ED3E7DBCDA9077 CRC64;
 Query Match 55.3%; Score 42; DB 2; Length 837;
 Best Local Similarity 58.3%; Pred. NO. 1.7e+02;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLRPSRVHNSSE 13
 |||:|:|
 DB 270 LLKPAYVHNEES 281

RESULT 18
 ID Q7NYT4 CHRVO PRELIMINARY; PRT; 451 AA.
 AC Q7NYT4;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Probable competence protein.


```

RL Science 307:1463-1465(2005) .
DR ENBL; CR626927; CAH05804.1; -; Genomic_DNA.
DR InterPro; IPR003702; ActCoA_hydro.
DR Pfam; PF02550; AcetylCoA_hydro; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 499 AA; 54064 MW; EB6F72A2B8B39295 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 499;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEE 13
DB 328 LILRPSEINSPE 340
:||||:||||

RESULT 20
Q650P9_BACFR
ID Q650P9_BACFR PRELIMINARY; PRT; 499 AA.
AC Q650P9; 2004 (TrEMBLrel. 28, Created)
DT 23-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative coenzyme A transferase.
GN OrderedLocusNames=BF0026;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OC NCBI_TaxID=817;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=YCH46;
RX PubMed=15466707; DOI=10.1073/pnas.0404172101;
RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA Kuhara S., Hattori M., Hayashi T., Ohnishi Y.;
RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT inversions regulating cell surface adaptation.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14919-14924(2004).
DR ENBL; AP006841; BAD46775.1; -; Genomic_DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006084; P:acetyl-CoA metabolism; IEA.
DR InterPro; IPR003702; ActCoA_hydro.
DR Pfam; PF02550; AcetylCoA_hydro; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 499 AA; 54063 MW; EB6F72A2B8B39295 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 499;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEE 13
DB 328 LILRPSEINSPE 340
:||||:||||

RESULT 21
Q57VF6_9TRYP
ID Q57VF6_9TRYP PRELIMINARY; PRT; 585 AA.
AC Q57VF6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Tb927.5.680;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OC NCBI_TaxID=5691;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
RA Shallow J., Hou L., Dijkens A., Feldblyum T., Hostettler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,

```

RA Khalak H.G., Salberg S., Simpson A.J., Tallon L., Van Aken S.,
 RL Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
 RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 KW EMBL; AC159441; AAX70413.1; -; Genomic_DNA.
 SQ SEQUENCE 585 AA; 65263 MW; 86BBA31F082073E5 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 585;
 Best Local Similarity 61.5%; Pred. No. 1.7e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
 :||| ||| |||
 Db 23 LLRRKRVHKEEN 35

RESULT 22
 Q9XUM0 CABEL PRELIMINARY; PRT; 740 AA.

AC Q9XUM0;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein W04A8.6.
 GN ORFNames=W04A8.6;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology."
 RL Science 282:2012-2018 (1998).

CC -!- INTERACTION:
 CC Q97238; gel-4; NExp=1; IntAct=EBI-326898, EBI-329192;
 CC Q22174; T04H1.2; NExp=1; IntAct=EBI-326898, EBI-311938;
 CC Q22227; T05C12.6a; NExp=1; IntAct=EBI-326898, EBI-316403;
 DR EMBL; Z82069; CAB04906.1; -; Genomic_DNA.
 DR PIR; T26140; T26140.
 DR IntAct; Q9XUM0; -.

DR Ensembl; W04A8.6; Caenorhabditis elegans.
 DR WormBase; WBGene0012241; W04A8.6.
 DR WormPeP; W04A8.6; CE16540.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 740 AA; 84732 MW; C0363C58BE7EDDE2 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 740;
 Best Local Similarity 46.2%; Pred. No. 2.2e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
 :||| ||| |||
 Db 204 MFKSKKNNKEN 216

RESULT 23

PHSG PASMU STANDARD; PRT; 818 AA.
 AC Q9CN90;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Glycogen phosphorylase (EC 2.4.1.1).
 GN Name=91gp; OrderedLocustNames=PW0545;
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
 CC -!- FUNCTION: Phosphorylase is an important allosteric enzyme in
 CC carbohydrate metabolism. Enzymes from different sources differ in
 CC their regulatory mechanisms and in their natural substrates.
 CC However, all known phosphorylases share catalytic and structural
 CC properties (By similarity).
 CC -!- CATALYTIC ACTIVITY: (1,4-alpha-D-glucosyl) (n) + phosphate = (1,4-
 CC alpha-D-glucosyl) (n-1) + alpha-D-glucose 1-phosphate.
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SIMILARITY: Belongs to the glycogen phosphorylase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

EMBL; AE006090; AAK02629.1; -; Genomic_DNA.

DR HSP; P06737; IEM6.
 DR InterPro; IPR011833; Glycophosphorylas.
 DR PANTHER; PTHR11468; Glyco_trans_35.
 DR Pfam; PF00343; Phosphorylase; 1.
 DR PIRSF; PIRSF00460; Ppylas; Glp; 1.
 DR TIGRFAMs; TIGR02093; P_ylas; 1.
 DR PROSITE; PS00102; PHOSPHORYLASE; 1.
 KW Allosteric enzyme; Carbohydrate metabolism; Complete proteome;
 KW Glycogen metabolism; Glycosyltransferase; Pyridoxal phosphate;
 KW Transferase.
 FT BINDING 567 667 Pyridoxal phosphate (covalent) (By
 FT similarity).

SQ SEQUENCE 818 AA; 94037 MW; 89E7643EDD48D5B1 CRC64;

Query Match 53.9%; Score 41; DB 1; Length 818;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEN 14
 :||| ||| |||
 Db 302 IYRKRHNTMEN 315

RESULT 24

Q9ZSB6 ARATH PRELIMINARY; PRT; 887 AA.
 AC Q9ZSB6;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE F3H7.7 protein.
 GN Name=F3H7.7;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
 OX NCBI_TaxID=3702;
 RN [1]

RP NUCLEOTIDE SEQUENCE.

RA WashU;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Laplant Y., Spalding L.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

[3]
 RN NUCLEOTIDE SEQUENCE.

RP Waterston R.;
 RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; AF118222; AAD03434.1; -; Genomic_DNA.
DR HSP; Q93009; INB.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001394; Peptidase C19.
DR InterPro; IPR006615; Pept_C19_N_1.
DR InterPro; IPR010460; Pept_C19_N_2.
DR Pfam; PF06337; DUF1055; 1.
DR Pfam; PF00443; UCH; 1.
DR SMART; SM00695; DUSP; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS00973; UCH_2_3; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
SQ SEQUENCE 887 AA; 99994 MW; 308BGFAB29C0B54 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 887;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHSEEN 14
Db 666 PSKIHGSEN 675

RESULT 25
Q93Y01 ARATH PRELIMINARY; PRT; 910 AA.
AC Q93Y01;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein At4g10590; T4F9.50 (At4g10590).
GN Name=At4g10590/T4F9.50; Synonym=At4g10590;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OC NCBI_TaxID=3702;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054463; AAK96655.1; -; mRNA.
DR EMBL; BT010363; AAK96806.1; -; mRNA.
DR HSP; Q93009; INB.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001394; Peptidase C19.
DR InterPro; IPR010460; Pept_C19_N_2.
DR Pfam; PF06337; DUF1055; 1.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
Hypothetical protein.

SQ SEQUENCE 910 AA; 102648 MW; E01B0AD928A8EE71 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 910;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHSEEN 14
Db 639 PSKIHGSEN 648

RESULT 26
Q9TOB8 ARATH PRELIMINARY; PRT; 937 AA.
AC Q9TOB8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein T4F9.50 (Hypothetical protein AT4g10590).
GN Name=T4F9.50; Synonym=At4g10590;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OC NCBI_TaxID=3702;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RP Bevan M., Wedler H., Wedler E., Wambutt R., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP EU Arabidopsis sequencing project;
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RP Wedler H., Wedler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP EMBL; AL161517; CAB78182.1; -; Genomic DNA.
DR PR; T04194; T04194.
DR HSP; Q93009; INB.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001394; Peptidase C19.
DR InterPro; IPR006615; Pept_C19_N_1.
DR InterPro; IPR010460; Pept_C19_N_2.
DR Pfam; PF06337; DUF1055; 1.
DR Pfam; PF00443; UCH; 1.
DR SMART; SM00695; DUSP; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
KW Hypothetical protein.

SQ SEQUENCE 937 AA; 105983 MW; 93E7A60C62EF4A4C CRC64;

Query Match 53.9%; Score 41; DB 2; Length 937;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PSRVHSEEN 14
Db 666 PSKIHGSEN 675

RESULT 27
Q9NR99_HUMAN
```

```

ID Q9NR99 HUMAN PRELIMINARY; PRT; 2828 AA.
AC Q9NR99;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Adican.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RA Crowl R.M., Luk D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245505; AAF86402.1; -, mRNA.
DR HSP; Q9B2K6; IOZN.
DR Ensembl; ENSG00000101825; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005021; F:vascular endothelial growth factor receptor. .; IEA.
DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR009134; VEGFR.
DR Pfam; PF00047; IG_4.
DR Pfam; PF00560; LRR_1_6.
DR PRINTS; PR01832; VEGFRECEPTOR.
DR SMART; SM00408; IGC2; 10.
DR SMART; SM00369; LRR_Typ; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00835; IG LIKE; 12.
DR Immunoglobulin domain; Leucine-rich repeat; Repeat.
KW SEQUENCE 2828 AA; 312294 MW; A18377D8554F1FE1 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 2828;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEN 14
:|:|:|:|:|:|
Db 2238 VVMKPAKIEHKEEN 2251

RESULT 28
Q8NFCO CORGL PRELIMINARY; PRT; 115 AA.
AC Q8NFCO; Q8M4C5;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein Cg11894.
GN OrderedLocNames=Cg11894, cg2074;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,

```

```

RA Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25 (2003).
DR EMBL; BA000036; BAB99287.1; -, Genomic DNA.
DR EMBL; BX927153; CAF20234.1; -, Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 115 AA; 12094 MW; 6B804C4BB455ED70 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 115;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PSRVHNSSEN 15
|:|:|:|:|:|
Db 43 PLRVHASSENT 53

RESULT 29
Q5CZC9 BACFN PRELIMINARY; PRT; 143 AA.
AC Q5CZC9;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Putative integral membrane protein.
GN OrderedLocNames=pbF9343.40c;
OS Bacteroides fragilis (strain ATCC 25285 / NCTC 9343).
OC Plasmid pbF9343.
OC Bacteria; Bacteroidetes; Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=272559;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15746427; DOI=10.1126/science.1107008;
RA Cerdeno-Tarraga A.-M., Patrick S., Crossman L.C., Blakely G.,
RA Abratt V., Lennard N., Foxton I., Duerden B., Harris B., Quail M.A.,
RA Barron A., Clark L., Corton C., Doggett J., Holden M.T.G., Larke N.,
RA Line A., Lord A., Norbertczak H., Ormond D., Price C.,
RA Rabinowitsch E., Woodward J., Barrell B.G., Parkhill J.;
RT "Extensive DNA inversions in the B. fragilis genome control variable
RT gene expression.";
RL Science 307:1463-1465 (2005).
DR EMBL; CR626928; CAH05754.1; -, Genomic DNA.
KW Complete proteome; Plasmid.
SQ SEQUENCE 143 AA; 16767 MW; 26DB2019C4DC6F03 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 143;
Best Local Similarity 55.6%; Pred. No. 47;
Matches 10; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 1 IVLRPSR---VHNSSEN 14
|:|:|:|:|:|:|:|:|
Db 70 IVMRPSRKAMQVFNEKN 87

RESULT 30
Q8ELX1 OCEIH PRELIMINARY; PRT; 174 AA.
AC Q8ELX1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein OB3093.
GN OrderedLocNames=OB3093;
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;

```

```

[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; BA00028; BAC15049.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 174 AA; 20245 MW; C56966C4B5A4DEC5 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 174;
Best Local Similarity 87.5%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VHNSEENT 15
Db 148 VHNNEENT 155
|||||
|::|::|

RESULT 31
Q6TV68_9POXV PRELIMINARY; PRT; 269 AA.
AC Q6TV68;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=129727;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=BV-AR02; DOI=10.1128/JVI.78.1.168-177.2004;
RX PubMed=14671098; PubMed=14671098; DOI=10.1128/JVI.78.1.168-177.2004;
RA Delhon G., Tulman E.R., Atonso C.L., Lu Z.,
RA de la Concha-Bermejillo A., Lehmkuhl H.D., Piccone M.E., Kutish G.F.,
RA Rock D.L.;
RT "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT Stomatitis Virus.";
RL J. Virol. 78:168-177(2004).
DR EMBL; AY386265; AAR98477.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 269 AA; 30567 MW; CC09C56720186501 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 269;
Best Local Similarity 56.7%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEENT 15
Db 49 RPSRQHSRSEENT 60
|||||
|::|::|

RESULT 32
Q6ZR62_HUMAN PRELIMINARY; PRT; 310 AA.
AC Q6ZR62;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46608.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
SQ TISSUE=Kidney;

[1]
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Maehuo Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128465; BAC87453.1; -; mRNA.
DR GO; GO:0003676; Pnucleic acid binding; IEA.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PROSITE; PS50158; ZF_CCHC; 1.
SQ SEQUENCE 310 AA; 34669 MW; 6C8D18EAA7E9A2A0 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 310;
Best Local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEENT 15
Db 23 ILRLPQMQHPTTENT 37
|||||
|::|::|

RESULT 33
Q6HWM1_BACHK PRELIMINARY; PRT; 350 AA.
AC Q6HWM1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein yheC.
GN Names=yheC; OrderedLocNames=BT9727_0766;
OS Bacillus thuringiensis (subsp. konkukian).
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=180856;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=97-27;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RT "Complete genome sequence of Bacillus thuringiensis 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT62687.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 350 AA; 40704 MW; 2D55A7256C6972C2 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 350;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 12
Db 145 IIRPTNHSSE 156
|||||
|::|::|

RESULT 34
Q81UK7_BACAN PRELIMINARY; PRT; 350 AA.
AC Q81UK7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=BA0861;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1392;
RN [1]
```

```

Query Match          52.6%;   Score 40;   DB 2;   Length 350;
Best Local Similarity 50.0%;   Pred. No. 1.4e+02;
Matches 6;   Conservative 4;   Mismatches 2;   Indels 0;   Gaps 0;

QY      1  IVLRPSRVHNS 12
      ||::||:|:|
DB      145  IIIRPTNHSSE 156

RESULT 36
Q4H676_9DEIO
ID      Q4H676_9DEIO PRELIMINARY;      PRT;      414 AA.
AC      Q4H676;
DT      13-SEP-2005 (TEMBLrel. 31, Created)
DT      13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DE      Hypothetical protein precursor.
GN      OFENAMES=DgeODRAFT 2770;
OS      Deinococcus geothermalis DSM 11300.
OC      Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC      Deinococcaceae; Deinococcus.
OX      NCBI_TaxID=319795;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=DSM 11300;
RG      US DOE Joint Genome Institute (JGI-PGF);
RA      Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA      Hammon N., Istrani S., Pitluck S., Richardson P.;
RT      "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT      DSM 11300."
RL      Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=DSM 11300;
RG      US DOE Joint Genome Institute (JGI-ORNL);
RA      Larimer F., Land M.;
RT      "Annotation of the draft genome assembly of Deinococcus geothermalis
RT      DSM 11300."
RL      Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC      -!- CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
CC      EMBL; AAHE01000016; EAL81802.1; -; Genomic_DNA.
DR      Hypothetical protein; Signal.
FW      SIGNAL
KW      1 29 Potential.
SQ      SEQUENCE 414 AA; 45121 MW; C13BB6981624FD9A CRC64;

Query Match          52.6%;   Score 40;   DB 2;   Length 414;
Best Local Similarity 58.3%;   Pred. No. 1.7e+02;
Matches 7;   Conservative 3;   Mismatches 2;   Indels 0;   Gaps 0;

QY      2  VLPRPSRVHNS 13
      .|:|:|:|:|
DB      363  VIIRTRVHPSQ 374

RESULT 37
FBX39 HUMAN
ID      FBX39 HUMAN STANDARD;      PRT;      442 AA.
AC      Q8N4B4;
DT      25-OCT-2004 (Rel. 45, Created)
DT      25-OCT-2004 (Rel. 45, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      F-box only protein 39.
GN      Name=FBXO39; Synonyms=FBX39;
OS      Homo sapiens (Human)
OC      Eukaryota; Metazoa; Chordata; Cranista; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC      Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

```

RC TISSUR=Brain;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
CC -1- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-
CC box protein)-type E3 ubiquitin ligase complex (By similarity).
CC -1- SUBUNIT: Directly interacts with SKP1A and CUL1 (By similarity).
CC -1- SIMILARITY: Contains 1 F-box domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BC034782; AAH34782.1; -; mRNA.
CC Ensembl; ENSG00000177294; Homo sapiens.
DR DR HGNC; HGNC:28565; FBXO39.
DR DR MIM; 609106; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00256; FBXO; 1.
DR PROSITE; PS00181; FBXO; FALSE_NEG.
DR Ub1 conjugation pathway.
KW DOMAIN 16 61 F-box.
FT SEQUENCE 442 AA; 52646 MW; 32A1E4D69FA29C71 CRC64;
SQ
Query Match 52.6%; Score 40; DB 1; Length 442;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPSRVHSE 12
Db ||||| ||
66 RPSRVHASE 74
RESULT 38
ID FBX39_MOUSE STANDARD; PRT; 443 AA.
AC Q5NBU5;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT F-box only protein 39.
DE F-box only protein 39.
GN Name=FBXO39;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG The mouse genome sequencing consortium;
Query Match 52.6%; Score 40; DB 1; Length 442;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPSRVHSE 12
Db ||||| ||
66 RPSRVHASE 74
RESULT 39
ID FBX39_RAT STANDARD; PRT; 443 AA.
AC Q56H10;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE F-box only protein 39.
GN Name=FBXO39;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1-431.
RC TISSUR=Testis;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE OF 343-443.
RP STRAIN=Sprague-Dawley;
RC Bonaldo M.F., Lennon G., Soares M.B.;
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-
CC box protein)-type E3 ubiquitin ligase complex (By similarity).
CC -1- SUBUNIT: Directly interacts with SKP1A and CUL1 (By similarity).
CC -1- SIMILARITY: Contains 1 F-box domain.
CC -1- CAUTION: This is a conceptual translation.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BC082089; AAH82089.1; ALT TERM; mRNA.
CC Ensembl; BF542717; -; NOT_ANNOTATED_CDS; mRNA.
DR Ensembl; ENSRNOG0000014953; Rattus norvegicus.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001611; LRR.

RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-
CC box protein)-type E3 ubiquitin ligase complex (By similarity).
CC -1- SUBUNIT: Directly interacts with SKP1A and CUL1 (By similarity).
CC -1- SIMILARITY: Contains 1 F-box domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL929071; CAI36035.1; ALT_INIT; Genomic DNA.
DR MGI; MGI:3505735; Fbxo39.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00181; FBXO; 1.
KW Ub1 conjugation pathway.
FT DOMAIN 13 59 F-box.
SQ SEQUENCE 443 AA; 52554 MW; 4F7D3B5F8B579213 CRC64;
Query Match 52.6%; Score 40; DB 1; Length 443;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 RPSRVHSE 12
Db ||||| ||
66 RPSRVHASE 74
RESULT 39
ID FBX39_RAT STANDARD; PRT; 443 AA.
AC Q56H10;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE F-box only protein 39.
GN Name=FBXO39;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1-431.
RC TISSUR=Testis;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE OF 343-443.
RP STRAIN=Sprague-Dawley;
RC Bonaldo M.F., Lennon G., Soares M.B.;
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-
CC box protein)-type E3 ubiquitin ligase complex (By similarity).
CC -1- SUBUNIT: Directly interacts with SKP1A and CUL1 (By similarity).
CC -1- SIMILARITY: Contains 1 F-box domain.
CC -1- CAUTION: This is a conceptual translation.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BC082089; AAH82089.1; ALT TERM; mRNA.
CC Ensembl; BF542717; -; NOT_ANNOTATED_CDS; mRNA.
DR Ensembl; ENSRNOG0000014953; Rattus norvegicus.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001611; LRR.

DR Pfam; PF00646; F-box; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR PROSITE; PS00181; FBOX; 1.
 KW Ubl conjugation pathway.
 FT DOMAIN 13 59 F-box.
 SQ SEQUENCE 443 AA; 52504 MW; 0A164BAC55B4B97F CRC64;
 Query Match 52.6%; Score 40; DB 1; Length 443;
 Best Local Similarity 88.9%; Pred. No. 1.8e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 4 RPSRVHNS 12
 Db 66 RPSRVHASE 74
 RESULT 40
 Q6G155 BARQU
 ID Q6G155 BARQU PRELIMINARY; PRT; 491 AA.
 AC Q6G155;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (G6PD).
 GN Name=zwf; OrderedLocustNames=BQ02930;
 OS Bartonella quintana (Rochalimaea quintana).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Bartonellaceae; Bartonella.
 OX NCBI_TaxID=803;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Toulouse;
 RX PubMed15210978; DOI=10.1073/pnas.0305659101;
 RA Almark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
 RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
 RA La Scola B., Holmberg M., Andersson S.G.E.;
 RT "The louse-borne human pathogen Bartonella quintana is a genomic
 RT derivative of the zoonotic agent Bartonella henselae";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).
 DR EMBL; BX897700; CAF25793.1; -; Genomic DNA.
 DR GO; GO:0004345; F:glucose-6-phosphate 1-dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006006; P:glucose metabolism; IEA.
 DR InterPro; IPR001282; G6PD.
 DR Pfam; PF02781; G6PD_C; 1.
 DR Pfam; PF00479; G6PD_N; 1.
 DR PIRSF; PIRSF000110; G6PD; 1.
 DR PRINTS; PR00079; G6PDHGRNASE.
 DR ProDom; PD001129; G6PD; 1.
 DR TIGRFAMs; TIGR00871; zwf; 1.
 DR PROSITE; PS00069; G6P_DSHYDROGNASE; 1.
 KW Carbohydrate metabolism; Complete proteome; Glucose metabolism; NADP;
 OX Oxidoreductase.
 SQ SEQUENCE 491 AA; 56166 MW; 4F7F3A4E9E8E107 CRC64;
 Query Match 52.6%; Score 40; DB 2; Length 491;
 Best Local Similarity 61.5%; Pred. No. 2.1e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 3 LRPSRVHNS 15
 Db 274 LMPNVHNSVEHT 286
 RESULT 41
 Q8L775 ARATH
 ID Q8L775 ARATH PRELIMINARY; PRT; 520 AA.
 AC Q8L775;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE RING finger-like protein (At5g24870).
 OX Arabidopsi

GN Name=At5g24870;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
 RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
 RA Deng J.M., Hayashizaki Y., Huan V.W., Lee J.M., Ishida J., Kamiya A.,
 RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,
 RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
 RA Theologis A., Davis R.W.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 DR EMBL; AY136427; AAM97093.1; -; mRNA.
 DR EMBL; BT014966; AAT47817.1; -; mRNA.
 DR HSRF; Q9LRB7; IlyM.
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 520 AA; 56792 MW; 6BBEC00519AC886F CRC64;
 Query Match 52.6%; Score 40; DB 2; Length 520;
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 IVLPRSVHNS 14
 Db 170 LVSRDARVNS 183
 RESULT 42
 CDPK DAUCA
 ID CDPK DAUCA STANDARD; PRT; 532 AA.
 AC P28582;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Calcium-dependent protein kinase (EC 2.7.1.-) (CDPK).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC asterids; campanulids; Apiales; Apiaceae; Scandiceae;
 OC Daucinae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Choi J.H.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 108-532.
 RX MEDLINE=92003674; PubMed=1912486;
 RA Suen K.-L., Choi J.H.;
 RT "Isolation and sequence analysis of a cDNA clone for a carrot calcium-
 RT dependent protein kinase: homology to calcium/calmodulin-dependent
 RT protein kinases and to calmodulin."
 RL Plant Mol. Biol. 17:581-590(1991).
 CC -1- FUNCTION: May play a role in signal transduction pathways that
 CC involve calcium as a second messenger.


```

CC CC -1- ENZYME REGULATION: Activated by calcium. Autophosphorylation may
CC CC play an important role in the regulation of the kinase activity.
CC CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. CamK
CC CC subfamily.
CC CC -1- SIMILARITY: Contains 4 EF-hand domains.
CC CC -----
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use as long as its content is in no way modified and this statement is not
CC CC removed.
CC CC -----
CC CC EMBL; X56599; CA3936.1; -; mRNA.
CC CC DR PIR; S17759; S17759.
CC CC DR PIR; T14335; T14335.
CC CC DR HSSP; Q63450; 1A06.
CC CC DR InterPro; IPR011992; EF-Hand type.
CC CC DR InterPro; IPR02048; EF-hand_Ca_bd.
CC CC DR InterPro; IPR00719; Prot_kinase.
CC CC DR InterPro; IPR008271; Ser_Thr_pkin_AS.
CC CC DR InterPro; IPR002290; Ser_Thr_pkinase.
CC CC DR Pfam; PF00036; ehand; 4.
CC CC DR Pfam; PF00069; Pkinase; 1.
CC CC DR ProDom; PD000012; EF-hand; 2.
CC CC DR ProDom; PD000001; Prot_kinase; 1.
CC CC DR SMART; SM00054; EFh; 4.
CC CC DR SMART; SM00220; S_TKC; 1.
CC CC DR PROSITE; PS00018; EF_HAND_1; 4.
CC CC DR PROSITE; PS00222; EF_HAND_2; 4.
CC CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC CC DR ATP-binding; Calcium; Kinase; Nucleotide-binding; Phosphorylation;
CC CC Repeat; Serine/threonine-protein kinase; Transferase.
CC CC KW ATP-binding; Calcium; Kinase; Nucleotide-binding; Phosphorylation;
CC CC FT DOMAIN 81 339 protein kinase.
CC CC FT DOMAIN 382 417 EF-hand 1.
CC CC FT DOMAIN 418 453 EF-hand 2.
CC CC FT DOMAIN 452 486 EF-hand 3.
CC CC FT DOMAIN 487 522 EF-hand 4.
CC CC FT NP_BIND 87 95 ATP (By similarity).
CC CC FT CA_BIND 395 406 1 (Potential).
CC CC FT CA_BIND 431 442 2 (Potential).
CC CC FT CA_BIND 465 476 3 (Potential).
CC CC FT CA_BIND 500 511 4 (Potential).
CC CC FT ACT_SITE 205 205 Proton acceptor (By similarity).
CC CC FT BINDING 110 110 ATP (By similarity).
CC CC SQ SEQUENCE 532 AA; 60065 MW; F03E6F036A0AE348 CRC64;

Query Match 52.6%; Score 40; DB 1; Length 532;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSENT 15
Db 56 LKPRQVHRPESNT 68

RESULT 43
ID Q5GZ92_XANOR PRELIMINARY; PRT; 771 AA.
AC Q5GZ92;
DC 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=XO2725;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;

```

```

RX PubMed=15673718; DOI=10.1093/nar/gki206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.-
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice."
RL Nucleic Acids Res 33:577-586(2005).
DR EMBL; AS013598; AAW75979.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR000160; GGDEF.
DR InterPro; IPR000144; PAS.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00563; EAL; 1.
DR Pfam; PF00990; GGDEF; 1.
DR Pfam; PF00072; Response_reg; 1.
DR SMART; SM00267; DUF1; 1.
DR SMART; SM00052; DUF2; 1.
DR TIGRFAMs; TIGR00254; GGDEF; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS50883; EAL; 1.
DR PROSITE; PS50887; GGDEF; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 771 AA; 85588 MW; B45DB7E9B918F9BD CRC64;

Query Match 52.6%; Score 40; DB 2; Length 771;
Best Local Similarity 61.5%; Pred. No. 3.5e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 13
Db 115 IAVRPSRPQNQEE 127

RESULT 44
ID Q6TQF0_XANCV PRELIMINARY; PRT; 834 AA.
AC Q6TQF0;
DC 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type III effector XopC.
GN Name=xopC;
OS Xanthomonas campestris (pv. vesicatoria).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=341;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=75-3;
RC PubMed=14645268; DOI=10.1128/JB.185.24.7092-7102.2003;
RA Noel L., Thieme F., Gabler J., Buttner D., Bonas U.;
RT "XopC and XopJ, two novel type III effector proteins from Xanthomonas
RT campestris pv. vesicatoria."
RL J. Bacteriol. 185:7092-7102(2003).
DR EMBL; AY389509; AAR23832.1; -; Genomic_DNA.
SQ SEQUENCE 834 AA; 91900 MW; 0A9D313AE315949E CRC64;

Query Match 52.6%; Score 40; DB 2; Length 834;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSENT 15
Db 220 LLRPTRDHYAEDST 233

RESULT 45
UM66_YEAST
ID UM66_YEAST STANDARD; PRT; 836 AA.

```

AC P39001;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transcriptional regulator UME6 (Negative transcriptional regulator of
DE IME2).
GN Name=UME6; Synonyms=CAR80, CARGRI, NIM2; OrderedLocusNames=YDR207C;
GN ORFNames=YD8142.04C;
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sigma 1278B;
RA Smart W.C., Park H.-D., Cooper T.G.;
RT "Sequence of the UME6/CAR80 gene from Saccharomyces cerevisiae";
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95011581; PubMed=7926768;
RA Strich R., Surosky R.T., Steber C., Messenguy F., Dubois E.,
RA Easton Esposito R.;
RT "UME6 is a key regulator of nitrogen repression and meiotic
RT development";
RL Genes Dev. 8:796-810(1994).
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S288c / GRF88;
RA Kumeno A.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN (4)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c / AB972;
RX MEDLINE=97313263; PubMed=9169867;
RA Jacq C., Alt-Moerbe J., Andre B., Arnold W., Ballesta J.P.G.,
RA Barques M., Baron L., Becker A., Biteau N., Bloeker H., Blugeon C.,
RA Boskovic J., Brandt P., Brueckner M., Buitrago M.J., Coster F.,
RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
RA Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Hankeln T.,
RA Hoheisel J.D., Jaeger P., Jimenez A., Jonniaux J.-L., Kraemer C.,
RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
RA Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
RA Paulin L., Perea J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M.A., Revuelta J.L.,
RA Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
RA Scharfe M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
RA Soler-Mira A., Urrestarazu L.A., Verhasselt P., Vissers S., Voet M.,
RA Volckaert G., Wagner G., Wambutt R., Wedler E., Wedler H., Woelfl S.,
RA Harris D.E., Bowman S., Brown D., Churcher C.M., Connor R., Dedman K.,
RA Gentles S., Hamlin N., Hunt S., Jones L., McDonald S., Murphy L.D.,
RA Niblett D., Odell C., Oliver K., Rajandream M.A., Richards C.,
RA Shore L., Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T.,
RA Allen E., Araujo R., Aviles E., Berno A., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Hunkle-Smith S., Hyman R.W.,
RA Komp C., Lashkari D., Le H., Lin D., Moseedale D., Nakahara K.,
RA Namath A., Oefner P., Oh C., Patel F.X., Roberts D., Schramm S.,
RA Schroeder M., Shogren T., Shroff N., Winant A., Yelton M.A.,
RA Botstein D., Davis R.W., Johnston M., Andrews S., Brinkman R.,
RA Cooper J., Ding H., Du Z., Favallo A., Fulton L., Gattung S.,
RA Greco T., Hallsworth K., Hawkins J., Hillier L.W., Jier M.,
RA Johnson D., Johnston L., Kirsten J., Kucaba T., Langston Y.,
RA Latreille P., Le T., Mardis E., Menezes S., Miller N., Nhan M.,
RA Pauley A., Peluso D., Rifkin L., Riles L., Taich A., Trevisan E.,
RA Vignati D., Wilcox L., Wohlman P., Vaudin M., Wilson R.,
RA Waterston R., Albermann K., Hani J., Heumann K., Kleine K.,
RA Newes H.-W., Zollner A., Zaccaria P.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV";
RL Nature 387:75-78(1997).
RN (5)
RP CHARACTERIZATION.
RX MEDLINE=92253412; PubMed=1579492;
RA Park H.-D., Luche R.M., Cooper T.G.;

RT "The yeast UME6 gene product is required for transcriptional
RT repression mediated by the CAR1 URS1 repressor binding site";
RL Nucleic Acids Res. 20:1909-1915(1992).
RN (6)
RP CHARACTERIZATION.
RX MEDLINE=96100650; PubMed=8528081;
RA Anderson S.F., Steber C.M., Easton Esposito R., Coleman J.E.;
RT "UME6, a negative regulator of meiosis in Saccharomyces cerevisiae,
RT contains a C-terminal Zn2Cys6 binuclear cluster that binds the URS1
RT DNA sequence in a zinc-dependent manner";
RL Protein Sci. 4:1832-1843(1995).
RN (7)
RP CHARACTERIZATION.
RX MEDLINE=9618868; PubMed=8614637; DOI=10.1093/nar/24.7.1322;
RA Jackson J.C., Lopes J.M.;
RT "The yeast UME6 gene is required for both negative and positive
RT transcriptional regulation of phospholipid biosynthetic gene
RT expression";
RL Nucleic Acids Res. 24:1322-1329(1996).
RN (8)
RP FUNCTION: Binds to the URS1 site (5'-AGCCGCCGA-3') and negatively
RP regulates the expression of many genes including CAR1 (arginase),
RP several required for sporulation, mating type switching, inositol
RP metabolism, and oxidative carbon metabolism. Has both a positive
RP and negative role in regulating phospholipid biosynthesis.
RN (9)
RP SUBUNIT: Interacts with IME1.
RN (10)
RP SUBCELLULAR LOCATION: Nuclear.
RN (11)
RP SIMILARITY: Contains 1 Zn(2)-C6 fungal-type DNA-binding domain.
RN (12)
RP This Swiss-Prot entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use as long as its content is in no way modified and this statement is not
RP removed.
RN (13)
RP EMBL; L32186; AAA34471.1; -; Genomic DNA.
RP EMBL; L24539; AAC14472.1; -; Unassigned DNA.
RP EMBL; L23663; BAA04890.1; -; Genomic DNA.
RP EMBL; Z68194; CAA92346.1; -; Genomic DNA.
RP PIR; S61570; S61570.
RP HSPF; P07272; IPI.
RP Interact; P39001; -;
RP GermOnline; 140699; -;
RP TransFam; T01247; -;
RP Ensembl; YDR207C; Saccharomyces cerevisiae.
RP SGD; S00002615; UME6.
RP GO; GO:0005634; C:nucleus; IPI.
RP GO; GO:0003677; F:DNA binding; IDA.
RP GO; GO:0030528; F:transcription regulator activity; IDA.
RP GO; GO:0007001; P:chromosome organization and biogenesis (sen. .; IMP.
RP GO; GO:0007131; P:meiotic recombination; IMP.
RP GO; GO:0007088; P:negative regulation of transcription, mitotic; IMP.
RP GO; GO:0045836; P:positive regulation of meiosis; IMP.
RP GO; GO:0030437; P:sporulation (sensu Fungi); IMP.
RP InterPro; IPR001138; Fungi_Tscrp_N.
RP Pfam; PF00172; Zn_C1us; 1.
RP SMART; SM00066; GAL4; 1.
RP PROSITE; PS00463; ZN2_Cy6_FUNGAL_1; 1.
RP PROSITE; PS00048; ZN2_Cy6_FUNGAL_2; 1.
KW Activator; Complete proteome; DNA-binding; Metal-binding;
KW Nuclear protein; Repressor; Transcription; Transcription regulation;
KW Zinc.
FT DNA BIND 771 798 Zn(2)-C6 fungal-type.
FT CONFLICT 101 101 V -> G (in Ref. 3).
FT CONFLICT 363 363 I -> V (in Ref. 1).
FT CONFLICT 443 443 N -> T (in Ref. 1).
FT CONFLICT 465 465 G -> D (in Ref. 1).
SQ SEQUENCE 836 AA; 91124 MW; 0DDA0A6B4A157182 CRC64;
Query Match 52.6%; Score 40; DB 1; Length 836;
Best Local Similarity 64.3%; Pred. No. 3.9e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
QY 2 VLRP--SRVHNSSE 13

```
Db 571 VLRPILLRIHNSQ 584
||||| |::|||:
CSE1_DROME STANDARD; PRT; 975 AA.
ID _CSE1_DROME
AC Q9XZU1; Q9UB14; Q9VJH4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Importin-alpha re-exporter (Cellular apoptosis susceptibility protein
DE homolog).
GN Name=Cas; ORFNames=CG13281;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21942515; PubMed=11944946; DOI=10.1006/dbio.2002.0612;
RA Tekotte H., Berndik D., Toerok T., Buszczak M., Jones L.M.,
RA Cooley L., Knoblich J.A., Davis I.;
RA "Dcas is required for importin-alpha nuclear export and mechano-
RA sensory organ cell fate specification in Drosophila.";
RL Dev. Biol. 244:396-406(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkley; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
```

```
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Berkley; TISSUE=Embryo;
RC MEDLINE=20196012; PubMed=10731138; DOI=10.1126/science.287.5461.2222;
RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
RA Stapleton M., Harvey D.A.;
RA "A Drosophila complementary DNA resource.";
RL Science 287:2222-2224(2000).
CC -!- FUNCTION: Export receptor for importin alpha. Mediates importin-
CC alpha reexport from the nucleus to the cytoplasm after import
CC substrates have been released into the nucleoplasm (By
CC similarity).
CC -!- SUBUNIT: Binds with high affinity to importin-alpha only in the
CC presence of RanGTP (By similarity).
CC -!- INTERACTION:
CC P23023-1:dxs; NtExp=1; IntAct=EBI-126291, EBI-196732;
CC Q9VX88:Gcn2; NtExp=1; IntAct=EBI-126291, EBI-137454;
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the CSE1 family.
CC -!- SIMILARITY: Contains 1 importin N-terminal domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ238857; CAB42967.1; -; mRNA..
CC ENBL; AS003652; AAF53575.1; -; Genomic_DNA.
CC EMBL; AF132562; AAD27861.1; -; mRNA..
CC IntAct; Q9XZU1; -.
CC Ensembl; CG13281; Drosophila melanogaster.
CC FlyBase; FBgn0022213; Cas.
CC GO; GO:0006611; P:protein-nucleus export; IMP.
CC InterPro; IPR011989; ARM-like.
CC InterPro; IPR005043; CAS_CSE1_C.
CC InterPro; IPR001494; Importinb_N.
CC Pfam; PF03378; CAS_CSE1; 1.
CC Pfam; PF03810; IBN_N; 1.
CC PROSITE; PS50166; IMPORTIN_B_NT; 1.
CC Nuclear protein; Protein transport; Transport.
CC DOMAIN 29 105 Importin N-terminal.
CC CONFLICT 194 194 K -> T (in Ref. 2).
CC CONFLICT 331 331 Q -> P (in Ref. 1).
CC CONFLICT 589 590 AA -> GR (in Ref. 1).
CC SEQUENCE 975 AA; 110168 MW; 5C9BA11C08D49C11 CRC64;
Query Match 52.6%; Score 40; DB 1; Length 975;
Best Local Similarity 54.5%; Pred. No. 4.7e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 5 PSRVHNSSEENT 15
Db 84 PDRIHESDRNT 94
RESULT 47
Q811Y8_PLAF7
ID Q811Y8_PLAF7 PRELIMINARY; PRT; 2249 AA.
AC Q811Y8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein PF00200C.
```

Mon Mar 13 10:54:35 2006

```
GN Name=PF0200c;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Barriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Fellwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagals K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seager K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.; 3-9 and 13."
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL: AL034557; CAD49133.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 2249 AA; 271088 MW; 5EF04691C22FB787 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 2249;
Best Local Similarity 54.5%; Pred. No. 1.3e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 PSRVHNSSEENT 15
DB 1799 PPKINNSEENT 1809

RESULT 48
Q7UWZ8 RHOB A PRELIMINARY; PRT; 7716 AA.
AC Q7UWZ8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL: BX294135; CAD72214.1; -; Genomic_DNA.
DR HSSP: P08661; 1B4.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0005529; F:sugar binding; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR003368; Chlamydia PMP.
DR InterPro: IPR002860; Glyco hydrolase.
DR InterPro: IPR001791; Laminin G.
```

```
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR006626; Pbh1.
DR InterPro: IPR001759; Pentaxin.
DR InterPro: IPR003129; TSP_N.
DR Pfam: PF02012; BNR; 5.
DR Pfam: PF00028; Cadherin; 6.
DR Pfam: PF02415; Chlam PMP; 7.
DR Pfam: PF02210; Laminin_G; 2; 2.
DR Pfam: PF00059; Lectin_C; 2.
DR Pfam: PF00354; Pentaxin; 1.
DR PRINTS: PR00205; CADHERIN.
DR PRINTS: PR00895; CADHERIN.
DR ProDom: PD002153; Pentaxin; 2.
DR SMART: SM00112; CA; 7.
DR SMART: SM00034; CLECT; 2.
DR SMART: SM00282; LamG; 3.
DR SMART: SM00710; Pbh1; 17.
DR TIGRFAMs: TIGR01376; POMP repeat; 3.
DR PROSITE: PS50268; CADHERIN_2; 8.
DR PROSITE: PS50041; C_TYPE LECTIN_2; 2.
DR PROSITE: PS50025; LAM_G DOMAIN; 1.
DR PROSITE: PS50025; LAM_G DOMAIN; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 7716 AA; 797868 MW; D391A25BD96405C0 CRC64;

Query Match 52.6%; Score 40; DB 2; Length 7716;
Best Local Similarity 63.6%; Pred. No. 5.5e+03;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 PSRVHNSSEENT 15
DB 153 PLRVHNSDDT 163

RESULT 49
QSRDW7 PONPY PRELIMINARY; PRT; 113 AA.
AC QSRDW7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459G0119.
GN Name=DKFZp459G0119;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbauer S., Schapp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR857776; CAH90040.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 13377 MW; CF0F91823777D34F CRC64;

Query Match 51.3%; Score 39; DB 2; Length 113;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RPSRVHNSSE 13
DB 48 RPDVHNSSE 57

RESULT 50
QSR4S3 PONPY PRELIMINARY; PRT; 139 AA.
AC QSR4S3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
```

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein DKFZp459F2016.
 GN Name=DKFZp459F2016;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
 CC Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Cortex;
 RG The German cDNA Consortium;
 RA Kohrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
 RA Pobo G., Han M., Wiemann S.;
 RA Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 RL EMBL; CR861171; CAH93243.1; -; mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 139 AA; 16177 MW; 5137F6EBA9B9AC05 CRC64;

 Query Match 51.3%; Score 39; DB 2; Length 139;
 Best Local Similarity 70.0%; Pred. No. 70;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 Qy 4 RPSRVHSEE 13
 . |||||
 Db 77 RPDVHNEE 86

 RESULT 51
 RT32 HUMAN
 ID RT32 HUMAN STANDARD; PRT; 142 AA.
 AC Q9F6G3; Q96Q48; Q9P0S1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Mitochondrial 28S ribosomal protein S32 (S32mt) (MRP-S32).
 GN Name=MRP32; Synonyms=MRP31, MRP42; ORFNames=HSPC204, PTD007;
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
 CC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Pituitary tumor;
 RA Mao Y.F., Peng Y., Dai M., Huang Q.H., Song H., Zhang Q.H., Mao M.,
 RA Fu G., Luo M., Chen J.H., Hu R.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Brain;
 RX MEDLINE=1154917; PubMed=11230166; DOI=10.1101/gr.154701;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glaesl S.,
 RA Anorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RX PubMed=14702039; DOI=10.1038/pg1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yanazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Moriyama Y., Momiya H., Sato N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Senba T.,
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Motomura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,
 RA Taguchi T., Oyama M., Hata H., Watanabe M., Komatsu T., Nakagawa K.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC Chen J.H., Luo W.Q., Hu S.N., Li G.T., Jin J., Huang X.W., Zhou H.J.,
 RA Yuan J.G., Qiang B.Q.;
 RT "Isolating a new human cDNA.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Umbilical cord blood;
 RX MEDLINE=20499367; PubMed=11042152; DOI=10.1101/gr.140200;
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 300 previously undefined genes expressed in CD34+ hematopoietic
 stem/progenitor cells.";
 RL Genome Res. 10:1546-1560(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Deje J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Litalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP NUCLEOTIDE SEQUENCE OF 78-128.
 RX MEDLINE=21429115; PubMed=11543634; DOI=10.1006/geno.2001.6622;
 RA Kenmochi N., Suzuki T., Uechi T., Magoori M., Kuniba M., Higa S.,
 RA Watanabe K., Tanaka T.;
 RT "The human mitochondrial ribosomal protein genes: mapping of 54 genes
 to the chromosomes and implications for human disorders.";
 RL Genomics 77:65-70(2001).
 RN [8]
 RP IDENTIFICATION.
 RX MEDLINE=21276436; PubMed=11279123; DOI=10.1074/jbc.M100727200;

```

RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Spremulli L.L.;
RT "The small subunit of the mammalian mitochondrial ribosome:
RL J. Biol. Chem. 276:19363-19374(2001).
CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- CAUTION: Was incorrectly assigned to be mitochondrial ribosomal
CC protein L42.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AF078860; AAD44492.1; -; mRNA.
DR EMBL; AF135160; AAF43784.1; -; mRNA.
DR EMBL; AK000285; BAA91054.1; -; mRNA.
DR EMBL; AL136659; CAB66594.1; -; mRNA.
DR EMBL; AF151038; AAF36124.1; -; mRNA.
DR EMBL; BC040240; AAH40240.1; -; mRNA.
DR EMBL; AB051626; BAB54953.1; -; Genomic DNA.
DR EMBL; ENSG00000198015; Homo sapiens.
DR HGNC; HGNC:14493; MRPL42.
DR GO; GO:0005763; C:mitochondrial small ribosomal subunit; IDA.
DR GO; GO:0003735; F:structural constituent of ribosome; NAS.
DR GO; GO:0006412; P:protein biosynthesis; NAS.
KW Mitochondrion; Ribonucleoprotein; Ribonucleoprotein.
SQ SEQUENCE 142 AA; 16661 MW; AED9ECFA653F870A CRC64;

Query Match 51.3%; Score 39; DB 1; Length 142;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy : 4 RPSRVHNSSE 13
|||:|
Db 77 RPDVHNNEE 86

RESULT 52
Q6FID1 HUMAN
ID Q6FID1_HUMAN PRELIMINARY; PRT; 142 AA.
AC Q6FID1
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE MRPL42 protein.
GN Name=MRPL42;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR533495; CAG38526.1; -; mRNA.
SQ SEQUENCE 142 AA; 16661 MW; AED9ECFA653F870A CRC64;

Query Match 51.3%; Score 39; DB 2; Length 142;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSE 13
|||:|
Db 77 RPDVHNNEE 86

RESULT 53
Q4R999 MACFA
ID Q4R999_MACFA PRELIMINARY; PRT; 212 AA.

```

```

AC Q4R999; 2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: Qtsa-10457, similar to human hypothetical protein
DE FLJ20291 (FLJ20291), (Crab eating macaque) (Cynomolgus monkey).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RT "DNA sequences of macaque genes expressed in brain or testis and its
RL evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB168197; BAF00322.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 212 AA; 24772 MW; 175A63CB14E13C2A CRC64;

Query Match 51.3%; Score 39; DB 2; Length 212;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSE 14
|||:|
Db 84 RPSKPHNSKVN 94

RESULT 54
Q4RF61 TETNG
ID Q4RF61_TETNG PRELIMINARY; PRT; 217 AA.
AC Q4RF61
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 14 SCAF15120, whole genome shotgun sequence.
DE (Fragment).
GN ORNames=GSTENG0035469001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percomorphia; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anghonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., de Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;

```



```

RL Submitted (PEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01015120; CAG12971.1; -; Genomic_DNA.
FT NON TER 217
SQ SEQUENCE 217 AA; 24829 MW; 4C59A115BE2D1DB0 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 217;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 15
Db 88 LRWSTLHGSE 100
|||:|||||
|||:|||||

RESULT 55
O6AL13 DESPS PRELIMINARY; PRT; 226 AA.
ID Q6AL13 DESPS PRELIMINARY; PRT; 226 AA.
AC Q6AL13;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=DP2063;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914; DOI=10.1111/J.1462-2920.2004.00665.x;
RA Rabus R., Ruepp A., Fricke T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments."
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG36792.1; -; Genomic DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 226 AA; 24666 MW; 492206D2FD96A9BA CRC64;

Query Match 51.3%; Score 39; DB 2; Length 226;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
Db 39 LMNSRLHNSSE 49
|||:|||||
|||:|||||

RESULT 56
Q4PI75 USTMA PRELIMINARY; PRT; 294 AA.
ID Q4PI75 USTMA PRELIMINARY; PRT; 294 AA.
AC Q4PI75;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM00188.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=521;
RC Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA

```

```

RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Bouhgalter B., Brunache A., Butler J., Caliste N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Doolley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseelis M., Karlsson E.,
RA Kells C., Kieu A., Kisaner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schuppach R., Seaman C., Settillali S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis."
RT Submitted (PEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACP01000002; EAK81573.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 294 AA; 32409 MW; 5B1A709E2BEBEC920 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 294;
Best Local Similarity 46.2%; Pred. No. 1.7e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLPSRVHNSSE 14
Db 137 ILRTSRINNTQN 149
|||:|||||
|||:|||||

RESULT 57
Q4WQC4 ASPFU PRELIMINARY; PRT; 316 AA.
ID Q4WQC4 ASPFU PRELIMINARY; PRT; 316 AA.
AC Q4WQC4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Alkaline phytoceramidase, putative.
GN ORFNames=Afu4g12330;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,

```

RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S., Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R., Fokker N., Fraser A., Garcia J.L., Garcia M.J., Goble A., Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B., Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J., Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R., Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C., Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M., Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I., Penvalva M.A., Pertea M., Price C., Pritchard B.L., Quail M.A., Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U., Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M., Ronning C.M., Rutter S., Salzberg S.L., Sanchez M., Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S., Takeuchi M., Tekai F., Turner G., Vazquez de Aldana C.R., Weidman J., White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K., Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR EMBL: AAHP01000005; EAL89560.1; -; Genomic DNA.
SQ SEQUENCE 316 AA; 37064 MW; F8824E1C3F42BB86 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 316;
Best Local Similarity 50.0%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 3;

Qy 1 IIVLRPSRVHNSSEN 14
Db 165 VTLRPSLRHSTED 178
:|||||:|:|:
:

RESULT 58
Q73G36 WOLPDM
ID Q73G36 WOLPDM PRELIMINARY; PRT; 316 AA.
AC Q73G36
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NADH-ubiquinone oxidoreductase, putative.
GN OrderedLocNames=WD1129;
OS Wolbachia pipiensis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T., Brownlie J.C., McGraw E.A., Martin W., Esser C., Almadinejad N., Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C., Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P., Berry K.J., Young M.B., Uterback T.R., Weidman J.F., Nierman W.C., Paulsen I.F., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipiensis wMel: a streamlined genome overrun by mobile genetic elements";
RL PLOS Biol. 2:327-341(2004).
DR EMBL: AS017260; AAS14782.1; -; Genomic DNA.
DR TIGR: WD1129;
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
KW Complete proteome; Ubiquinone.
SQ SEQUENCE 316 AA; 35718 MW; B3D773BEA6BFE326 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 316;
Best Local Similarity 50.0%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 3; Mismatches 4;

Qy 1 IIVLRPSRVHNSSEN 14
Db 165 VTLRPSLRHSTED 178
:|||||:|:|:
:

RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R., Fokker N., Fraser A., Garcia J.L., Garcia M.J., Goble A., Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B., Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J., Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R., Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C., Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M., Mouyna I., Mulligan S., Murphy L., O'Neill S., Paulsen I., Penvalva M.A., Pertea M., Price C., Pritchard B.L., Quail M.A., Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U., Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M., Ronning C.M., Rutter S., Salzberg S.L., Sanchez M., Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S., Takeuchi M., Tekai F., Turner G., Vazquez de Aldana C.R., Weidman J., White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K., Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR EMBL: AAHP01000005; EAL89560.1; -; Genomic DNA.
SQ SEQUENCE 316 AA; 37064 MW; F8824E1C3F42BB86 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 316;
Best Local Similarity 50.0%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 3;

Qy 1 IIVLRPSRVHNSSEN 14
Db 165 VTLRPSLRHSTED 178
:|||||:|:|:
:

RESULT 58
Q73G36 WOLPDM
ID Q73G36 WOLPDM PRELIMINARY; PRT; 316 AA.
AC Q73G36
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NADH-ubiquinone oxidoreductase, putative.
GN OrderedLocNames=WD1129;
OS Wolbachia pipiensis wMel.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=66077;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15024419; DOI=10.1371/journal.pbio.0020069;
RA Wu M., Sun L.V., Vamathevan J.J., Riegler M., DeBoy R.T., Brownlie J.C., McGraw E.A., Martin W., Esser C., Almadinejad N., Wiegand C., Madupu R., Beanan M.J., Brinkac L.M., Daugherty S.C., Durkin A.S., Kolonay J.F., Nelson W.C., Mohamoud Y., Lee P., Berry K.J., Young M.B., Uterback T.R., Weidman J.F., Nierman W.C., Paulsen I.F., Nelson K.E., Tettelin H., O'Neill S.L., Eisen J.A.;
RT "Phylogenomics of the reproductive parasite Wolbachia pipiensis wMel: a streamlined genome overrun by mobile genetic elements";
RL PLOS Biol. 2:327-341(2004).
DR EMBL: AS017260; AAS14782.1; -; Genomic DNA.
DR TIGR: WD1129;
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
KW Complete proteome; Ubiquinone.
SQ SEQUENCE 316 AA; 35718 MW; B3D773BEA6BFE326 CRC64;

Db 147 IIVRPSLVFGKEDN 160
RESULT 59
Q9MOV1 ARATH
ID Q9MOV1 ARATH PRELIMINARY; PRT; 333 AA.
AC Q9MOV1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE D123-like protein.
GN Name=AT4G05440;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M., Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M., Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161503; CAB81086.1; -; Genomic_DNA.
DR PIR: D85068; D85068.
DR InterPro: IPR009772; D123.
DR Pfam: PF07065; D123.
SQ SEQUENCE 333 AA; 38777 MW; 85DDBA4B4F38703D2 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 333;
Best Local Similarity 77.8%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 1;

Qy 5 PSRVHNSSE 13
Db 58 PSRVHNSSE 66
:|||||:
:

RESULT 60
Q6D5I9 ERWCT
ID Q6D5I9 ERWCT PRELIMINARY; PRT; 364 AA.
AC Q6D5I9
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative ABC-transporter ATP-binding protein.
GN OrderedLocNames=ECA2051;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J., Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K., Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H., Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S., Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia carotovora subsp. atroseptica and characterization of virulence factors";
RT Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
RL EMBL: EX950851; CAG74953.1; -; Genomic_DNA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.

DR	GO: 0016887; F:ATPase activity; IEA.
DR	GO: 0000166; F:nucleotide binding; IEA.
DR	GO: 0001547; F:polyamine-transporting ATPase activity; IEA.
DR	GO: 0015846; P:polyamine transport; IEA.
DR	InterPro: IPR003593; AAA ATPase.
DR	InterPro: IPR003439; ABC_transp.-like.
DR	InterPro: IPR005893; Sp/PrABC_transp.
DR	Pfam: PF00005; ABC_tran; 1.
DR	ProDom: PD000006; ABC transporter; 1.
DR	SMART: SM00382; AAA; 1.
DR	TIGRFAMs: TIGR01187; potA; 1.
DR	PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR	PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
DR	ATP-binding; Complete proteome.
DR	SEQUENCE 364 AA; 41307 MW; 676CE6C1F284ABF1 CRC64;

Query Match 51.3%; Score 39; DB 2; Length 364;
Best Local Similarity 40.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy	1	I	V	L	R	P	S	R	V	H	N	S	E	E	N	T	15
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	280	L	C	I	R	P	E	R	I	H	V	N	E	A	H	T	294

Search completed: March 11, 2006, 12:24:37
Job time : 169.842 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:25:33 ; Search time 33.5526 Seconds
(without alignments)
36.961 Million cell updates/sec

Title: US-10-774-242A-1
Perfect score: 76
Sequence: 1 IVLPRSRVHSEENT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCUTS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	547	2	US-09-949-016-8330 Sequence 8330, Ap
2	76	100.0	760	1	US-08-230-491A-2 Sequence 2, Appli
3	76	100.0	760	1	US-08-619-280A-2 Sequence 2, Appli
4	76	100.0	760	1	US-08-940-391-2 Sequence 2, Appli
5	76	100.0	760	2	US-09-268-606-2 Sequence 2, Appli
6	52	68.4	761	2	US-09-518-550-26 Sequence 26, Appli
7	42	55.3	327	2	US-09-248-786A-23053 Sequence 23053, A
8	40	52.6	836	2	US-09-538-092-173 Sequence 173, App
9	40	52.6	1584	2	US-09-251-645-6 Sequence 6, Appli
10	39	51.3	81	2	US-09-270-767-62451 Sequence 62451, A
11	39	51.3	144	2	US-09-673-395A-306 Sequence 306, App
12	39	51.3	144	2	US-09-673-395A-588 Sequence 588, App
13	39	51.3	389	2	US-09-270-767-46824 Sequence 46824, A
14	38	50.0	368	2	US-09-538-092-357 Sequence 357, App
15	38	50.0	371	2	US-09-212-247C-6 Sequence 6, Appli
16	38	50.0	371	2	US-10-076-157-6 Sequence 6, Appli
17	37.5	49.3	710	2	US-09-252-991A-20733 Sequence 20733, Ap
18	37	48.7	83	1	US-08-499-676A-12 Sequence 12, Appli
19	37	48.7	113	2	US-09-270-767-61152 Sequence 61152, A
20	37	48.7	293	2	US-09-270-767-37828 Sequence 37828, A
21	37	48.7	293	2	US-09-270-767-53045 Sequence 53045, A
22	37	48.7	366	2	US-09-830-433A-4 Sequence 4, Appli
23	37	48.7	375	2	US-09-489-039A-7670 Sequence 7670, Ap
24	37	48.7	414	2	US-09-489-039A-12578 Sequence 12578, A
25	37	48.7	414	1	US-07-816-283-10 Sequence 10, Appli
26	37	48.7	418	1	US-08-417-103-10 Sequence 10, Appli
27	37	48.7	418	2	US-09-826-509-573 Sequence 573, App

28	37	48.7	544	1	US-08-264-002-7 Sequence 7, Appli
29	37	48.7	544	2	US-09-457-040B-15 Sequence 15, Appli
30	37	48.7	1021	2	US-09-270-767-45640 Sequence 45640, A
31	36	47.4	62	2	US-09-252-991A-27405 Sequence 27405, A
32	36	47.4	172	2	US-09-248-796A-24931 Sequence 24931, A
33	36	47.4	191	2	US-09-270-767-33410 Sequence 33410, A
34	36	47.4	212	2	US-08-858-207A-381 Sequence 381, App
35	36	47.4	212	2	US-09-270-767-33440 Sequence 33440, A
36	36	47.4	212	2	US-09-270-767-48657 Sequence 48657, A
37	36	47.4	219	2	US-09-252-991A-29516 Sequence 29516, A
38	36	47.4	269	2	US-09-134-000C-5252 Sequence 5252, Ap
39	36	47.4	292	2	US-09-190-824-2 Sequence 2, Appli
40	36	47.4	295	2	US-09-583-110-5136 Sequence 5136, Ap
41	36	47.4	420	2	US-09-107-433-3370 Sequence 3370, Ap
42	36	47.4	420	2	US-09-213-053-4 Sequence 4, Appli
43	36	47.4	469	1	US-08-477-451-23 Sequence 23, Appli
44	36	47.4	501	2	US-09-934-901-20 Sequence 20, Appli
45	36	47.4	501	2	US-09-934-868-10 Sequence 10, Appli
46	36	47.4	501	2	US-10-321-210-20 Sequence 20, Appli
47	36	47.4	501	2	US-10-320-874-20 Sequence 20, Appli
48	36	47.4	501	2	US-10-320-924-20 Sequence 20, Appli
49	36	47.4	559	1	US-10-701-200-10 Sequence 10, Appli
50	36	47.4	559	1	US-08-313-553-15 Sequence 15, Appli
51	36	47.4	559	2	US-08-767-993-15 Sequence 15, Appli
52	36	47.4	645	2	US-09-540-236-2083 Sequence 2083, Ap
53	36	47.4	653	2	US-09-248-796A-20015 Sequence 20015, A
54	36	47.4	1089	2	US-09-252-991A-20334 Sequence 20334, A
55	36	47.4	1250	1	US-08-441-139-9 Sequence 9, Appli
56	36	47.4	1250	2	US-09-487-558B-364 Sequence 364, App
57	35.5	46.7	1394	2	US-09-213-053-2 Sequence 2, Appli
58	35	46.1	1467	2	US-09-252-991A-17657 Sequence 17657, A
59	35	46.1	72	2	US-09-513-999C-6876 Sequence 6876, Ap
60	35	46.1	73	2	US-09-471-276-1285 Sequence 1285, A
61	35	46.1	81	2	US-09-270-767-59565 Sequence 59565, A
62	35	46.1	83	1	US-08-499-676A-11 Sequence 10, Appli
63	35	46.1	83	1	US-08-439-676A-11 Sequence 11, Appli
64	35	46.1	91	2	US-09-252-991A-23940 Sequence 23940, A
65	35	46.1	100	2	US-09-270-767-38626 Sequence 38626, A
66	35	46.1	100	2	US-09-270-767-53843 Sequence 53843, A
67	35	46.1	189	2	US-09-270-767-45559 Sequence 45559, A
68	35	46.1	213	2	US-08-529-055-31 Sequence 31, Appli
69	35	46.1	271	2	US-09-270-767-44144 Sequence 44144, A
70	35	46.1	301	2	US-09-902-540-15018 Sequence 15018, A
71	35	46.1	336	2	US-09-636-215-590 Sequence 590, App
72	35	46.1	347	2	US-09-636-215-590 Sequence 590, App
73	35	46.1	347	2	US-09-685-166A-590 Sequence 590, App
74	35	46.1	347	2	US-09-679-426-590 Sequence 590, App
75	35	46.1	347	2	US-09-759-143-590 Sequence 590, App
76	35	46.1	347	2	US-09-651-236-590 Sequence 590, App
77	35	46.1	347	2	US-09-657-279-590 Sequence 590, App
78	35	46.1	347	2	US-10-012-896-590 Sequence 590, App
79	35	46.1	374	1	US-08-820-170A-25 Sequence 25, Appli
80	35	46.1	374	2	US-09-055-699-25 Sequence 25, Appli
81	35	46.1	374	2	US-09-273-565-25 Sequence 25, Appli
82	35	46.1	374	2	US-09-565-538-25 Sequence 25, Appli
83	35	46.1	374	2	US-09-661-468-25 Sequence 25, Appli
84	35	46.1	374	2	US-09-976-165-25 Sequence 25, Appli
85	35	46.1	374	2	US-09-227-853A-2 Sequence 2, Appli
86	35	46.1	381	2	PCT-US95-06385-2 Sequence 92, App
87	35	46.1	428	1	US-07-816-283-12 Sequence 12, Appli
88	35	46.1	428	1	US-08-417-103-12 Sequence 12, Appli
89	35	46.1	428	1	US-09-540-236-2267 Sequence 2267, Ap
90	35	46.1	514	2	US-09-489-039A-7647 Sequence 7647, Ap
91	35	46.1	538	2	US-09-248-796A-21009 Sequence 21009, A
92	35	46.1	685	2	US-09-541-094-18 Sequence 18, Appli
93	35	46.1	692	2	US-09-945-258-18 Sequence 18, Appli
94	35	46.1	695	2	US-09-252-991A-17440 Sequence 17440, A
95	35	46.1	713	2	US-09-824-574-3 Sequence 3, Appli
96	35	46.1	1703	2	US-09-487-558B-340 Sequence 340, App
97	35	46.1	1703	2	US-08-714-741-32 Sequence 32, Appli
98	35	46.1	8991	1	US-08-307-499-6 Sequence 6, Appli
99	34.5	45.4	114	1	US-09-299-268-6 Sequence 6, Appli
100	34.5	45.4	114	2	US-09-299-268-6 Sequence 6, Appli

101	34	44.7	52	2	US-09-755-665-42	Sequence 42, Appl	174	33	43.4	123	2	US-09-270-767-58078	Sequence 58078, A
102	34	44.7	81	2	US-09-248-796A-24938	Sequence 24938, A	175	33	43.4	151	2	US-09-270-767-465947	Sequence 45947, A
103	34	44.7	100	2	US-09-755-665-40	Sequence 40, Appl	176	33	43.4	156	2	US-09-270-767-46154	Sequence 46154, A
104	34	44.7	103	2	US-09-270-767-39273	Sequence 39273, A	177	33	43.4	163	2	US-09-248-796A-20018	Sequence 20018, A
105	34	44.7	103	2	US-09-270-767-54490	Sequence 54490, A	178	33	43.4	190	2	US-09-252-991A-23782	Sequence 23782, A
106	34	44.7	107	2	US-09-755-665-10	Sequence 10, Appl	179	33	43.4	192	2	US-09-248-796A-17782	Sequence 17782, A
107	34	44.7	117	2	US-09-893-737-208	Sequence 208, App	180	33	43.4	200	2	US-09-540-236-2527	Sequence 2527, Ap
108	34	44.7	118	2	US-09-902-540-16302	Sequence 16302, A	181	33	43.4	200	2	US-09-270-767-31918	Sequence 31918, A
109	34	44.7	158	2	US-08-849-634B-2	Sequence 2, Appl	182	33	43.4	211	2	US-09-248-796A-28127	Sequence 28127, A
110	34	44.7	160	2	US-08-902-540-10806	Sequence 10806, A	183	33	43.4	219	2	Sequence 12942, A	Sequence 12942, A
111	34	44.7	177	2	US-09-893-737-166	Sequence 166, App	184	33	43.4	241	2	Sequence 1894, Ap	Sequence 1894, Ap
112	34	44.7	181	2	US-09-107-532A-6013	Sequence 6013, Ap	185	33	43.4	253	4	Sequence 52, Appl	Sequence 52, Appl
113	34	44.7	195	2	US-08-849-634B-3	Sequence 3, Appl	186	33	43.4	258	2	Sequence 4764, Ap	Sequence 4764, Ap
114	34	44.7	224	2	US-09-976-594-187	Sequence 187, App	187	33	43.4	266	2	Sequence 114, App	Sequence 114, App
115	34	44.7	224	2	US-09-919-039-95	Sequence 95, Appl	188	33	43.4	310	2	Sequence 6082, Ap	Sequence 6082, Ap
116	34	44.7	238	2	US-09-270-767-31899	Sequence 31899, A	189	33	43.4	313	2	Sequence 43932, A	Sequence 43932, A
117	34	44.7	238	2	US-09-270-767-47116	Sequence 47116, A	190	33	43.4	320	1	Sequence 17, Appl	Sequence 17, Appl
118	34	44.7	239	2	US-09-252-991A-28204	Sequence 28204, A	191	33	43.4	332	2	Sequence 3771, Ap	Sequence 3771, Ap
119	34	44.7	257	2	US-09-543-681A-7481	Sequence 7481, Ap	192	33	43.4	340	2	Sequence 262, App	Sequence 262, App
120	34	44.7	289	2	US-09-489-039A-9173	Sequence 9173, Ap	193	33	43.4	374	2	Sequence 42761, A	Sequence 42761, A
121	34	44.7	312	2	US-09-270-767-57576	Sequence 57576, A	194	33	43.4	380	2	Sequence 33, Appl	Sequence 33, Appl
122	34	44.7	314	2	US-09-270-767-42737	Sequence 42737, A	195	33	43.4	437	2	Sequence 49, Appl	Sequence 49, Appl
123	34	44.7	315	2	US-09-370-398-6	Sequence 6, Appl	196	33	43.4	438	2	Sequence 30, Appl	Sequence 30, Appl
124	34	44.7	315	2	US-10-090-190-6	Sequence 6, Appl	197	33	43.4	438	2	Sequence 32, Appl	Sequence 32, Appl
125	34	44.7	315	2	US-10-082-902-6	Sequence 6, Appl	198	33	43.4	438	2	Sequence 34, Appl	Sequence 34, Appl
126	34	44.7	330	2	US-09-270-767-32670	Sequence 32670, A	199	33	43.4	438	2	Sequence 36, Appl	Sequence 36, Appl
127	34	44.7	330	2	US-09-270-767-47887	Sequence 47887, A	200	33	43.4	438	2	Sequence 38, Appl	Sequence 38, Appl
128	34	44.7	342	2	US-09-270-767-42293	Sequence 42293, A	201	33	43.4	438	2	Sequence 40, Appl	Sequence 40, Appl
129	34	44.7	350	2	US-09-489-039A-12799	Sequence 12799, A	202	33	43.4	438	2	Sequence 42, Appl	Sequence 42, Appl
130	34	44.7	370	2	US-09-543-681A-5868	Sequence 5868, Ap	203	33	43.4	438	2	Sequence 44, Appl	Sequence 44, Appl
131	34	44.7	416	2	US-09-489-039A-8918	Sequence 8918, Ap	204	33	43.4	438	2	Sequence 46, Appl	Sequence 46, Appl
132	34	44.7	445	2	US-09-328-352-4631	Sequence 4631, Ap	205	33	43.4	441	2	Sequence 26086, A	Sequence 26086, A
133	34	44.7	477	2	US-09-252-991A-31798	Sequence 31798, A	206	33	43.4	449	2	Sequence 35750, Ap	Sequence 35750, Ap
134	34	44.7	491	2	US-09-248-796A-18483	Sequence 18483, A	207	33	43.4	494	2	Sequence 6244, Ap	Sequence 6244, Ap
135	34	44.7	492	2	US-09-270-767-59114	Sequence 59114, A	208	33	43.4	507	2	Sequence 18770, A	Sequence 18770, A
136	34	44.7	588	2	US-09-710-279-3328	Sequence 3328, Ap	209	33	43.4	519	2	Sequence 4524, Ap	Sequence 4524, Ap
137	34	44.7	593	2	US-09-134-001C-3592	Sequence 3592, Ap	210	33	43.4	529	2	Sequence 12604, A	Sequence 12604, A
138	34	44.7	615	1	US-08-738-349-12	Sequence 12, Appl	211	33	43.4	532	2	Sequence 4071, Ap	Sequence 4071, Ap
139	34	44.7	693	1	US-08-738-349-6	Sequence 6, Appl	212	33	43.4	553	2	Sequence 32290, A	Sequence 32290, A
140	34	44.7	693	2	US-09-919-497-55	Sequence 55, Appl	213	33	43.4	559	2	Sequence 2, Appl	Sequence 2, Appl
141	34	44.7	728	2	US-09-298-924-2	Sequence 2, Appl	214	33	43.4	586	2	Sequence 2, Appl	Sequence 2, Appl
142	34	44.7	730	2	US-09-270-767-43721	Sequence 43721, A	215	33	43.4	628	2	Sequence 2844, Ap	Sequence 2844, Ap
143	34	44.7	748	1	US-08-937-931-4	Sequence 4, Appl	216	33	43.4	682	2	Sequence 17856, A	Sequence 17856, A
144	34	44.7	748	2	US-09-285-502-4	Sequence 4, Appl	217	33	43.4	755	2	Sequence 14733, A	Sequence 14733, A
145	34	44.7	748	2	US-09-709-126-4	Sequence 4, Appl	218	33	43.4	764	1	Sequence 2, Appl	Sequence 2, Appl
146	34	44.7	748	2	US-09-871-385A-4	Sequence 4, Appl	219	33	43.4	764	1	Sequence 325, App	Sequence 325, App
147	34	44.7	782	1	US-07-725-083-2	Sequence 2, Appl	220	33	43.4	764	1	Sequence 6, Appl	Sequence 6, Appl
148	34	44.7	782	2	US-08-669-286-10	Sequence 10, Appl	221	33	43.4	769	1	Sequence 11021, A	Sequence 11021, A
149	34	44.7	782	2	US-09-469-253-10	Sequence 10, Appl	222	33	43.4	769	1	Sequence 544, App	Sequence 544, App
150	34	44.7	782	2	US-09-642-146-10	Sequence 10, Appl	223	33	43.4	798	2	Sequence 10169, A	Sequence 10169, A
151	34	44.7	787	2	US-09-949-016-9446	Sequence 9446, Ap	224	33	43.4	798	2	Sequence 5, Appl	Sequence 5, Appl
152	34	44.7	796	1	US-08-189-228-58	Sequence 58, Appl	225	33	43.4	831	2	Sequence 25809, A	Sequence 25809, A
153	34	44.7	796	1	US-08-332-643-52	Sequence 52, Appl	226	33	43.4	834	2	Sequence 5284, Ap	Sequence 5284, Ap
154	34	44.7	796	1	US-08-332-638-58	Sequence 58, Appl	227	33	43.4	838	2	Sequence 69, Appl	Sequence 69, Appl
155	34	44.7	796	1	US-08-738-349-2	Sequence 2, Appl	228	33	43.4	900	2	Sequence 4, Appl	Sequence 4, Appl
156	34	44.7	796	1	US-08-738-349-4	Sequence 4, Appl	229	33	43.4	950	2	Sequence 11048, A	Sequence 11048, A
157	34	44.7	796	2	US-08-654-328-2	Sequence 2, Appl	230	33	43.4	1055	2	Sequence 20, Appl	Sequence 20, Appl
158	34	44.7	1300	2	US-10-102-549-2	Sequence 2, Appl	231	33	43.4	1075	2	Sequence 58, Appl	Sequence 58, Appl
159	34	44.7	1437	2	US-09-949-016-10535	Sequence 10535, A	232	33	43.4	1121	2	Sequence 1, Appl	Sequence 1, Appl
160	34	44.7	1457	2	US-08-665-259-27	Sequence 27, Appl	233	33	43.4	1344	2	Sequence 322, App	Sequence 322, App
161	34	44.7	1457	2	US-08-762-500-27	Sequence 27, Appl	234	33	43.4	1466	2	Sequence 4, Appl	Sequence 4, Appl
162	34	44.7	1472	2	US-09-032-438C-119	Sequence 119, App	235	33	43.4	1469	2	Sequence 11048, A	Sequence 11048, A
163	34	44.7	1980	2	US-09-914-272A-3	Sequence 3, Appl	236	33	43.4	1471	2	Sequence 2452, Ap	Sequence 2452, Ap
164	34	44.7	1980	2	US-10-638-333-3	Sequence 3, Appl	237	33	43.4	1492	2	Sequence 2, Appl	Sequence 2, Appl
165	34	44.7	1980	2	US-10-747-133A-3	Sequence 3, Appl	238	33	43.4	1536	2	Sequence 1, Appl	Sequence 1, Appl
166	34	44.7	2560	2	US-09-949-016-11554	Sequence 11554, A	239	33	43.4	1536	2	Sequence 322, App	Sequence 322, App
167	33.5	44.1	270	1	US-08-611-880-1	Sequence 1, Appl	240	33	43.4	1472	1	Sequence 4, Appl	Sequence 4, Appl
168	33.5	44.1	270	2	US-09-085-305-13	Sequence 13, Appl	241	33	43.4	4472	1	Sequence 2, Appl	Sequence 2, Appl
169	33	43.4	48	2	US-09-583-110-4391	Sequence 4391, Ap	242	32	42.1	5518	2	Sequence 274, App	Sequence 274, App
170	33	43.4	83	2	US-09-270-767-39816	Sequence 39816, A	243	32	42.1	42	2	Sequence 2054, A	Sequence 2054, A
171	33	43.4	83	2	US-09-270-767-55033	Sequence 55033, A	244	32	42.1	67	2	Sequence 6060, A	Sequence 6060, A
172	33	43.4	89	2	US-09-746-311B-115	Sequence 115, App	245	32	42.1	73	2	Sequence 387, Ap	Sequence 387, Ap
173	33	43.4	102	2	US-09-333-809-198	Sequence 198, App	246	32	42.1	105	2	Sequence 3817, Ap	Sequence 3817, Ap

247	32	42.1	125	2	US-09-370-838-59	Sequence 59, Appl	320	32	42.1	386	2	US-09-769-402-2	Sequence 2, Appl
248	32	42.1	125	2	US-09-854-133-59	Sequence 59, Appl	321	32	42.1	386	2	US-09-949-002-357	Sequence 357, App
249	32	42.1	126	2	US-09-270-767-46788	Sequence 46788, A	322	32	42.1	396	2	US-09-252-991A-31744	Sequence 31744, A
250	32	42.1	128	2	US-09-370-838-61	Sequence 61, Appl	323	32	42.1	399	2	US-09-949-002-437	Sequence 437, App
251	32	42.1	128	2	US-09-252-991A-28235	Sequence 28235, A	324	32	42.1	416	2	US-10-272-490-81	Sequence 81, Appl
252	32	42.1	128	2	US-09-854-133-61	Sequence 61, Appl	325	32	42.1	416	2	US-10-272-490-87	Sequence 87, Appl
253	32	42.1	132	2	US-09-270-767-33546	Sequence 33546, A	326	32	42.1	417	2	US-09-820-002-4	Sequence 4, Appl
254	32	42.1	132	2	US-09-270-767-48763	Sequence 48763, A	327	32	42.1	418	2	US-09-270-767-38036	Sequence 38036, A
255	32	42.1	149	2	US-09-270-767-32661	Sequence 32661, A	328	32	42.1	418	2	US-09-270-767-53253	Sequence 53253, A
256	32	42.1	149	2	US-09-270-767-47878	Sequence 47878, A	329	32	42.1	424	2	US-08-871-268A-23	Sequence 23, Appl
257	32	42.1	159	2	US-09-618-259-5	Sequence 5, Appl	330	32	42.1	424	2	US-08-871-267B-31	Sequence 31, Appl
258	32	42.1	163	2	US-09-489-039A-10183	Sequence 10183, A	331	32	42.1	424	2	US-09-618-419-31	Sequence 31, Appl
259	32	42.1	163	2	US-09-902-540-10547	Sequence 10547, A	332	32	42.1	424	2	US-09-163-674-23	Sequence 23, Appl
260	32	42.1	164	2	US-09-518-046-25	Sequence 25, Appl	333	32	42.1	435	2	US-09-784-810C-13	Sequence 13, Appl
261	32	42.1	164	2	US-09-650-371-25	Sequence 25, Appl	334	32	42.1	435	2	US-09-248-796A-15251	Sequence 15251, A
262	32	42.1	174	2	US-09-270-767-35941	Sequence 35941, A	335	32	42.1	445	2	US-10-272-490-2	Sequence 2, Appl
263	32	42.1	174	2	US-09-270-767-51158	Sequence 51158, A	336	32	42.1	450	2	US-09-949-016-7182	Sequence 7182, Ap
264	32	42.1	180	2	US-09-252-991A-32665	Sequence 32665, A	337	32	42.1	452	2	US-09-583-110-3905	Sequence 3905, Ap
265	32	42.1	183	2	US-09-312-283C-383	Sequence 383, App	338	32	42.1	455	2	US-10-272-490-34	Sequence 34, Appl
266	32	42.1	183	2	US-10-012-231A-68	Sequence 68, Appl	339	32	42.1	458	2	US-09-270-767-61004	Sequence 61004, A
267	32	42.1	183	2	US-10-015-389A-68	Sequence 68, Appl	340	32	42.1	477	2	US-09-107-433-3353	Sequence 3353, Ap
268	32	42.1	183	2	US-10-006-768A-68	Sequence 68, Appl	341	32	42.1	489	2	US-09-667-135-30	Sequence 30, Appl
269	32	42.1	183	2	US-10-015-671A-68	Sequence 68, Appl	342	32	42.1	492	2	US-09-792-024-91	Sequence 91, Appl
270	32	42.1	183	2	US-10-011-833A-68	Sequence 68, Appl	343	32	42.1	496	2	US-08-949-637-2	Sequence 2, Appl
271	32	42.1	183	2	US-10-011-833A-68	Sequence 68, Appl	344	32	42.1	496	2	US-09-291-488-2	Sequence 2, Appl
272	32	42.1	183	2	US-10-006-041A-68	Sequence 68, Appl	345	32	42.1	508	2	US-09-902-540-10388	Sequence 10388, A
273	32	42.1	183	2	US-10-012-064A-68	Sequence 68, Appl	346	32	42.1	509	2	US-09-902-540-11469	Sequence 11469, A
274	32	42.1	183	2	US-10-030-269A-2	Sequence 2, Appl	347	32	42.1	567	2	US-09-328-352-6937	Sequence 6937, Ap
275	32	42.1	197	1	US-08-989-289-1	Sequence 1, Appl	348	32	42.1	579	1	US-08-128-564A-31	Sequence 31, Appl
276	32	42.1	197	1	US-09-328-866-1	Sequence 1, Appl	349	32	42.1	579	4	PCT-US94-09143-31	Sequence 31, Appl
277	32	42.1	197	2	US-09-220-399-2	Sequence 2, Appl	350	32	42.1	601	2	US-09-270-767-45496	Sequence 45496, A
278	32	42.1	197	2	US-10-017-754-1904	Sequence 1904, Ap	351	32	42.1	633	2	US-09-538-092-786	Sequence 786, App
279	32	42.1	216	2	US-09-270-767-44613	Sequence 44613, A	352	32	42.1	640	2	US-09-252-991A-27604	Sequence 27604, A
280	32	42.1	218	2	US-09-489-039A-10326	Sequence 10326, A	353	32	42.1	658	2	US-09-252-991A-26319	Sequence 26319, A
281	32	42.1	222	2	US-10-101-464A-720	Sequence 720, App	354	32	42.1	661	2	US-09-371-338-7	Sequence 7, Appl
282	32	42.1	223	2	US-09-538-092-145	Sequence 145, App	355	32	42.1	687	2	US-09-107-433-3382	Sequence 3382, Ap
283	32	42.1	226	2	US-09-107-532A-3696	Sequence 3696, Ap	356	32	42.1	690	2	US-09-907-794A-49	Sequence 49, Appl
284	32	42.1	231	2	US-09-487-558B-232	Sequence 232, App	357	32	42.1	690	2	US-09-905-125A-49	Sequence 49, Appl
285	32	42.1	237	2	US-09-252-991A-21759	Sequence 21759, A	358	32	42.1	690	2	US-09-902-775A-49	Sequence 49, Appl
286	32	42.1	255	2	US-08-944-483-67	Sequence 67, Appl	359	32	42.1	690	2	US-09-906-700-49	Sequence 49, Appl
287	32	42.1	256	1	US-09-027-337-3	Sequence 3, Appl	360	32	42.1	690	2	US-09-903-603A-49	Sequence 49, Appl
288	32	42.1	256	2	US-09-644-600-3	Sequence 3, Appl	361	32	42.1	690	2	US-09-904-920A-49	Sequence 49, Appl
289	32	42.1	256	2	US-09-654-600A-3	Sequence 3, Appl	362	32	42.1	690	2	US-09-904-920A-49	Sequence 49, Appl
290	32	42.1	273	2	US-09-270-767-41185	Sequence 41185, A	363	32	42.1	690	2	US-09-909-064-49	Sequence 49, Appl
291	32	42.1	273	2	US-09-270-767-56401	Sequence 56401, A	364	32	42.1	690	2	US-09-905-381A-49	Sequence 49, Appl
292	32	42.1	280	1	US-08-312-387B-6	Sequence 6, Appl	365	32	42.1	690	2	US-09-906-618-49	Sequence 49, Appl
293	32	42.1	280	1	US-08-683-426-6	Sequence 6, Appl	366	32	42.1	690	2	US-09-906-646-49	Sequence 49, Appl
294	32	42.1	280	1	US-08-683-458-6	Sequence 6, Appl	367	32	42.1	690	2	US-09-904-462-49	Sequence 49, Appl
295	32	42.1	280	1	US-08-878-360-6	Sequence 6, Appl	368	32	42.1	690	2	US-09-902-736A-49	Sequence 49, Appl
296	32	42.1	280	2	US-08-478-140B-6	Sequence 6, Appl	369	32	42.1	694	2	US-09-538-092-1351	Sequence 1351, Ap
297	32	42.1	280	2	US-09-333-412-6	Sequence 6, Appl	370	32	42.1	698	2	US-09-949-016-10215	Sequence 10215, A
298	32	42.1	280	2	US-09-338-943-6	Sequence 6, Appl	371	32	42.1	701	2	US-09-583-110-4537	Sequence 4537, App
299	32	42.1	283	2	US-10-007-267A-6	Sequence 6, Appl	372	32	42.1	707	2	US-09-538-092-288	Sequence 288, App
300	32	42.1	283	2	US-09-252-991A-22824	Sequence 22824, A	373	32	42.1	726	6	5208144-37	Patent No. 5208144
301	32	42.1	296	2	US-09-270-767-39170	Sequence 39170, A	374	32	42.1	731	2	US-09-708-426-12	Sequence 12, Appl
302	32	42.1	296	2	US-09-270-767-54387	Sequence 54387, A	375	32	42.1	731	2	US-09-949-016-11729	Sequence 11729, A
303	32	42.1	299	2	US-09-389-956-16	Sequence 16, Appl	376	32	42.1	749	2	US-09-902-540-15370	Sequence 15370, A
304	32	42.1	311	2	US-09-270-767-42099	Sequence 42099, A	377	32	42.1	751	2	US-09-902-540-15370	Sequence 15370, A
305	32	42.1	317	2	US-09-746-801A-57	Sequence 57, Appl	378	32	42.1	885	1	US-08-484-106-14	Sequence 14, Appl
306	32	42.1	317	2	US-10-719-885-57	Sequence 57, Appl	379	32	42.1	885	1	US-08-484-106-14	Sequence 14, Appl
307	32	42.1	326	2	US-09-248-796A-16985	Sequence 16985, A	380	32	42.1	894	2	US-09-688-188B-18	Sequence 18, Appl
308	32	42.1	349	2	US-10-300-819B-25	Sequence 25, Appl	381	32	42.1	894	2	US-09-291-417D-18	Sequence 18, Appl
309	32	42.1	353	2	US-09-107-532A-7032	Sequence 7032, Ap	382	32	42.1	911	2	US-09-252-991A-19929	Sequence 19929, A
310	32	42.1	353	2	US-09-248-796A-20008	Sequence 20008, A	383	32	42.1	941	2	US-09-513-783A-172	Sequence 172, App
311	32	42.1	355	2	US-09-252-991A-32375	Sequence 32375, A	384	32	42.1	941	2	US-10-100-957A-172	Sequence 172, App
312	32	42.1	356	2	US-09-248-796A-15198	Sequence 15198, A	385	32	42.1	1002	2	US-10-290-579A-184	Sequence 184, App
313	32	42.1	370	2	US-09-252-991A-24608	Sequence 24608, A	386	32	42.1	1177	2	US-09-270-767-59152	Sequence 59152, A
314	32	42.1	370	2	US-08-311-731A-203	Sequence 203, App	387	32	42.1	1177	2	US-09-270-767-62395	Sequence 62395, A
315	32	42.1	376	2	US-09-820-002-2	Sequence 2, Appl	388	32	42.1	1673	2	US-09-418-710-70	Sequence 70, Appl
316	32	42.1	386	2	US-09-086-483A-2	Sequence 2, Appl	389	32	42.1	1673	2	US-09-839-479-69	Sequence 69, Appl
317	32	42.1	386	2	US-08-875-082-2	Sequence 2, Appl	390	32	42.1	1674	2	US-09-418-710-1	Sequence 1, Appl
318	32	42.1	386	2	US-09-130-491-6	Sequence 6, Appl	391	32	42.1	1674	2	US-09-839-479-1	Sequence 1, Appl
319	32	42.1	386	2	US-09-580-212-2	Sequence 2, Appl	392	32	42.1	1970	2	US-09-538-092-1005	Sequence 1005, Ap

393	32	42.1	2008	2	US-09-270-767-43754	Sequence 43754, A	466	31	40.8	242	2	US-09-270-767-58952	Sequence 58952, A
394	32	42.1	2008	2	US-09-270-767-46774	Sequence 46774, A	467	31	40.8	242	2	US-09-248-796A-14587	Sequence 14587, A
395	32	42.1	15281	1	US-08-471-119A-2	Sequence 2, Appl1	468	31	40.8	245	2	US-09-252-991A-29983	Sequence 29983, A
396	31.5	41.4	258	2	US-09-248-796A-18803	Sequence 18803, A	469	31	40.8	245	2	US-09-489-039A-8432	Sequence 8432, Ap
397	31.5	41.4	486	2	US-09-248-796A-18727	Sequence 18727, A	470	31	40.8	245	2	US-09-134-000C-6032	Sequence 6032, Ap
398	31.5	41.4	809	2	US-09-489-039A-18275	Sequence 18275, A	471	31	40.8	247	1	US-08-200-005C-2	Sequence 2, Appl1
399	31	40.8	16	1	US-07-942-245-142	Sequence 142, App	472	31	40.8	247	1	US-09-038-597A-2	Sequence 10, Appl1
400	31	40.8	16	1	US-07-942-245-177	Sequence 177, App	473	31	40.8	247	1	US-08-460-529B-10	Sequence 2, Appl1
401	31	40.8	16	1	US-08-808-130-4	Sequence 4, Appl1	474	31	40.8	247	1	US-08-431-117A-2	Sequence 10, Appl1
402	31	40.8	38	2	US-09-270-767-36792	Sequence 36792, A	475	31	40.8	247	2	US-09-361-736B-10	Sequence 10, Appl1
403	31	40.8	52	2	US-09-270-767-36792	Sequence 36792, A	476	31	40.8	247	2	US-09-949-016-6319	Sequence 6319, Ap
404	31	40.8	52	2	US-09-270-767-36792	Sequence 36792, A	477	31	40.8	247	2	US-09-787-397A-4	Sequence 4, Appl1
405	31	40.8	52	2	US-09-513-999C-6845	Sequence 6845, Ap	478	31	40.8	247	2	US-09-705-500A-3	Sequence 3, Appl1
406	31	40.8	66	2	US-09-513-999C-4407	Sequence 4407, Ap	479	31	40.8	247	2	US-10-418-226-10	Sequence 10, Appl1
407	31	40.8	74	2	US-09-248-796A-23340	Sequence 23340, A	480	31	40.8	249	2	US-09-134-001C-5030	Sequence 5030, Ap
408	31	40.8	80	1	US-08-848-252-2	Sequence 2, Appl1	481	31	40.8	262	2	US-09-252-991A-23161	Sequence 23161, A
409	31	40.8	83	2	US-09-621-976-4139	Sequence 4139, Ap	482	31	40.8	272	2	US-10-232-858-75	Sequence 75, Appl1
410	31	40.8	85	2	US-09-107-532A-4720	Sequence 4720, Ap	483	31	40.8	272	2	US-09-338-063A-75	Sequence 75, Appl1
411	31	40.8	90	2	US-09-270-767-41019	Sequence 41019, A	484	31	40.8	273	2	US-09-270-767-39703	Sequence 39703, A
412	31	40.8	90	2	US-09-270-767-56235	Sequence 56235, A	485	31	40.8	273	2	US-09-270-767-54920	Sequence 54920, A
413	31	40.8	101	2	US-09-732-210-250	Sequence 250, App	486	31	40.8	279	2	US-09-270-767-47699	Sequence 47699, A
414	31	40.8	101	2	US-09-248-796A-23591	Sequence 23591, Ap	487	31	40.8	293	2	US-09-896-096A-18	Sequence 18, Appl1
415	31	40.8	109	2	US-09-513-999C-6258	Sequence 6258, Ap	488	31	40.8	300	2	US-09-602-787A-198	Sequence 198, App
416	31	40.8	110	2	US-09-270-767-32810	Sequence 32810, A	489	31	40.8	302	2	US-09-248-796A-18445	Sequence 18445, A
417	31	40.8	118	2	US-09-270-767-48027	Sequence 48027, A	490	31	40.8	302	2	US-09-248-796A-20893	Sequence 20893, A
418	31	40.8	123	2	US-09-509-559B-1	Sequence 1, Appl1	491	31	40.8	303	2	US-09-328-352-7497	Sequence 7497, Ap
419	31	40.8	125	2	US-09-732-210-72	Sequence 72, Appl1	492	31	40.8	304	2	US-09-583-110-3507	Sequence 3507, Ap
420	31	40.8	125	2	US-09-732-210-100	Sequence 100, App	493	31	40.8	308	2	US-09-438-185A-834	Sequence 834, App
421	31	40.8	125	2	US-09-732-210-102	Sequence 102, App	494	31	40.8	318	2	US-09-248-796A-24411	Sequence 24411, A
422	31	40.8	125	2	US-09-732-210-272	Sequence 272, App	495	31	40.8	319	2	US-09-198-452A-893	Sequence 893, App
423	31	40.8	125	2	US-09-732-210-1695	Sequence 1695, Ap	496	31	40.8	320	2	US-09-347-803-23	Sequence 23, Appl1
424	31	40.8	125	2	US-09-732-210-1696	Sequence 1696, Ap	497	31	40.8	321	2	US-10-232-858-80	Sequence 80, Appl1
425	31	40.8	125	2	US-09-732-210-1697	Sequence 1697, Ap	498	31	40.8	321	2	US-09-338-063A-80	Sequence 80, Appl1
426	31	40.8	125	2	US-09-732-210-1708	Sequence 1708, Ap	499	31	40.8	323	2	US-09-270-767-55931	Sequence 55931, A
427	31	40.8	125	2	US-09-732-210-1709	Sequence 1709, Ap	500	31	40.8	323	2	US-09-270-767-51148	Sequence 51148, A
428	31	40.8	125	2	US-09-732-210-1710	Sequence 1710, Ap							
429	31	40.8	125	2	US-09-732-210-1715	Sequence 1715, Ap							
430	31	40.8	137	2	US-09-540-236-2262	Sequence 2262, Ap							
431	31	40.8	138	2	US-09-252-991A-20131	Sequence 20131, A							
432	31	40.8	151	2	US-10-104-047-3705	Sequence 3705, Ap							
433	31	40.8	153	2	US-09-252-991A-32543	Sequence 32543, A							
434	31	40.8	154	2	US-10-104-047-1992	Sequence 1992, Ap							
435	31	40.8	157	2	US-09-518-046-23	Sequence 23, Appl1							
436	31	40.8	157	2	US-09-650-371-23	Sequence 23, Appl1							
437	31	40.8	159	2	US-09-280-839-3	Sequence 3, Appl1							
438	31	40.8	166	2	US-09-081-180-4	Sequence 4, Appl1							
439	31	40.8	166	2	US-09-040-786-4	Sequence 9, Appl1							
440	31	40.8	168	2	US-08-444-628-9	Sequence 9, Appl1							
441	31	40.8	168	2	US-08-357-820-9	Sequence 9, Appl1							
442	31	40.8	170	2	US-09-732-210-1676	Sequence 1676, Ap							
443	31	40.8	170	2	US-09-902-540-15981	Sequence 15981, A							
444	31	40.8	171	1	US-08-303-270-1	Sequence 1, Appl1							
445	31	40.8	175	2	US-09-060-726A-2	Sequence 2, Appl1							
446	31	40.8	175	2	US-09-845-849A-2	Sequence 2, Appl1							
447	31	40.8	177	2	US-09-507-242-2	Sequence 2, Appl1							
448	31	40.8	180	2	US-09-248-796A-20046	Sequence 20046, A							
449	31	40.8	183	2	US-09-270-767-35830	Sequence 35830, A							
450	31	40.8	183	2	US-09-270-767-51047	Sequence 51047, A							
451	31	40.8	185	2	US-10-018-924-6	Sequence 6, Appl1							
452	31	40.8	193	2	US-08-655-352-2	Sequence 2, Appl1							
453	31	40.8	193	2	US-09-258-016-2	Sequence 2, Appl1							
454	31	40.8	193	2	US-09-257-825B-2	Sequence 2, Appl1							
455	31	40.8	200	2	US-09-252-991A-31738	Sequence 31738, A							
456	31	40.8	200	2	US-09-949-016-10310	Sequence 10310, A							
457	31	40.8	212	2	US-09-543-681A-6648	Sequence 6648, Ap							
458	31	40.8	222	2	US-09-270-767-40120	Sequence 40120, A							
459	31	40.8	222	2	US-09-270-767-55336	Sequence 55336, A							
460	31	40.8	227	2	US-09-902-540-10904	Sequence 10904, A							
461	31	40.8	231	2	US-09-134-000C-5664	Sequence 5664, Ap							
462	31	40.8	231	2	US-09-538-092-721	Sequence 721, App							
463	31	40.8	234	2	US-09-252-991A-25162	Sequence 25162, A							
464	31	40.8	234	2	US-09-270-767-34127	Sequence 34127, A							
465	31	40.8	234	2	US-09-270-767-49344	Sequence 49344, A							

ALIGNMENTS

RESULT 1
US-09-949-016-8330
; Sequence 8330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8330
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8330

Query Match 100.0%; Score 76; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IVLRPSRVHNSSENT 15
Db 31 IVLRPSRVHNSSENT 45

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-619-280A-2

Query Match 100.0%; Score 76; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSENT 15
Db 24 IVLPRSRVHNSSENT 38
|||||
|||||

RESULT 4
US-08-940-391-2
; Sequence 2, Application US/08940391
; Patent No. 5965373
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
; TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

RESULT 2
 US-08-230-491A-2
 Sequence 2, Application US/08230491A
 Patent No. 5587299
 GENERAL INFORMATION:
 APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
 APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
 TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FELPE & LYNCH
 STREET: 805 THIRD AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: WORDPERFECT - ASC II
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/230,491A
 FILING DATE: 20-APRIL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5587299man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 330
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 760 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-230-491A-2
 Query Match 100.0%; Score 76; DB 1; Length 760;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 IVLPRSRVHNSSEENT 15
 Db 24 IVLPRSRVHNSSEENT 38
 RESULT 3
 US-08-619-280A-2
 Sequence 2, Application US/08619280A
 Patent No. 5767242
 GENERAL INFORMATION:
 APPLICANT: Zimmermann, Rainer; Park, John E.;
 APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
 TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
 TITLE OF INVENTION: ALPHA, AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felpe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect

US-08-940-391-2

Query Match 100.0%; Score 76; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSENT 15
Db 24 IVLRPSRVHNSSENT 38

RESULT 5

US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. 6846910

; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265.606
; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-09-265-606-2

Query Match 100.0%; Score 76; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSENT 15
Db 24 IVLRPSRVHNSSENT 38

RESULT 6

US-09-518-550-26
; Sequence 26, Application US/09518550
; Patent No. 6875851

; GENERAL INFORMATION:
; APPLICANT: TRAVIS, James
; APPLICANT: FOTEMPA, Jan
; APPLICANT: BANBULA, Agnieszka

; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE

; FILE REFERENCE: 235.00190101
; CURRENT APPLICATION NUMBER: US/09/518,550
; CURRENT FILING DATE: 2000-03-03

; PRIOR APPLICATION NUMBER: 60/123,148
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US00/05551
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 26
; LENGTH: 761
; TYPE: PRT

; ORGANISM: Mus musculus
US-09-518-550-26

Query Match 68.4%; Score 52; DB 2; Length 761;
Best Local Similarity 73.3%; Pred. No. 0.43;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSENT 15
Db 24 IVLRPSRVKPEGNT 38

RESULT 7

US-09-248-796A-23053
; Sequence 23053, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS;
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 23053
; LENGTH: 327

; TYPE: PRT

; ORGANISM: Candida albicans

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (2),(3),(21)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-09-248-796A-23053

Query Match 55.3%; Score 42; DB 2; Length 327;
Best Local Similarity 53.8%; Pred. No. 1.1;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IVLRPSRVHNSSENT 14
Db 87 VLNPSQISNNEDN 99

RESULT 8

US-09-538-092-173
; Sequence 173, Application US/09538092
; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

```

; PRIOR APPLICATION NUMBER: 60/178,965
; FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 173
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number YDR207C
US-09-538-092-173

Query Match      52.6%; Score 40; DB 2; Length 836;
Best Local Similarity 64.3%; Pred. No. 77;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 2 VLRP--SRVHNSSEE 13
Db 571 VLRPILLRIHNSSEQ 584

RESULT 9
US-09-251-645-6
; Sequence 6, Application US/09251645
; Patent No. 6281413
; GENERAL INFORMATION:
; APPLICANT: Kramer, Vance C.
; APPLICANT: Morgan, Michael K.
; APPLICANT: Anderson, Arne R.
; APPLICANT: Hart, Hope
; APPLICANT: Warren, Gregory W.
; APPLICANT: Dunn, Martha
; APPLICANT: Chen, Jeng S.
; TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: CGC1963/A
; CURRENT APPLICATION NUMBER: US/09/251,645
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-09-251-645-6

Query Match      52.6%; Score 40; DB 2; Length 1584;
Best Local Similarity 61.5%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEE 13
Db 35 IIVLRPNVGNSEQ 47

RESULT 10
US-09-270-767-62451
; Sequence 62451, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62451
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

US-09-270-767-62451
Query Match      51.3%; Score 39; DB 2; Length 81;
Best Local Similarity 77.8%; Pred. No. 8.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 RVHNSSEENT 15
Db 26 RVHNIEDNT 34

RESULT 11
US-09-673-395A-306
; Sequence 306, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 306
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-306

Query Match      51.3%; Score 39; DB 2; Length 144;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEE 13
Db 79 RPDVHNSSEE 88

RESULT 12
US-09-673-395A-588
; Sequence 588, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 588
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-588

Query Match      51.3%; Score 39; DB 2; Length 144;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHNSSEE 13
Db 79 RPDVHNSSEE 88

```


RESULT 17
US-09-252-991A-20733
; Sequence 20733, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20733
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20733

Query Match 49.3%; Score 37.5; DB 2; Length 710;
Best Local Similarity 56.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 IVLRP-SRVHSEENT 15
Db 545 VVLRPLDEVHDEVT 560

RESULT 18
US-08-499-676A-12
; Sequence 12, Application US/08499676A
; Patent No. 5998154
; GENERAL INFORMATION:
; APPLICANT: AGNES SCHONBRUNN
; TITLE OF INVENTION: SOMATOSTATIN RECEPTOR PEPTIDE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/499,676A
; FILING DATE: July 7, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Krieger, Paul E.
; REGISTRATION NUMBER: 25,886
; REFERENCE/DOCKET NUMBER: 79247.3/A95175US
; TELEPHONE: 713-850-0909
; TELEFAX: 713-850-0165
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-499-676A-12

Query Match 48.7%; Score 37; DB 1; Length 83;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEENT 15
Db 16 VLLRPSRRVRSQPT 30

RESULT 19
US-09-270-767-61152
; Sequence 61152, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 61152
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61152

Query Match 48.7%; Score 37; DB 2; Length 113;
Best Local Similarity 46.2%; Pred. No. 28;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLRPSRVHSEEN 14
Db 9 LMDPTLVHTREN 21

RESULT 20
US-09-270-767-37828
; Sequence 37828, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 37828
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-37828

Query Match 48.7%; Score 37; DB 2; Length 293;
Best Local Similarity 33.3%; Pred. No. 82;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEENT 15
Db 203 VICNPKQMSQENS 217

RESULT 21
US-09-270-767-53045
; Sequence 53045, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

```

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53045
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-53045

```

Query Match 48.7%; Score 37; DB 2; Length 293;
Best Local Similarity 33.3%; Pred. No. 82;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEENT 15
pb 203 VICNPKOMNSOENS 217

RESULT 22
US-09-830-433A-4
: Sequence 4. Application US/09830433A

```

; Sequence 4, Application US/09830433A
;
; Patent No. 6835384
; GENERAL INFORMATION:
;
; APPLICANT: AUJAME et al.
; TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
; strains of the Neisseria genus
; FILE REFERENCE: P07180US00/BAS
; CURRENT APPLICATION NUMBER: US/09/830.433A
; CURRENT FILING DATE: 2001-04-26
;
; PRIOR APPLICATION NUMBER: FR 98 13 693
;

```

Query Match	48.7%	Score 37;	DB 2;	Length 366;
Best Local Similarity	54.5%	Pred. No. 1.1e+02;		
Matches	6:	Conservative	2:	Mismatches 3: Indels 0: Gaps 0;

Qy 4 RPSRVHNSEEN 14
||| :|| :|
Db 167 RPSRLHNLRON 177

```

RESULT 23
US-09-489-039A-7670
; Sequence 7670, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2094001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7670
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7670

```

Query Match 48.7%; Score 37; DB 2; Length 375;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels

Qy 1 IVLRPSRVHNSENT 15
:|:|:|:
Dp 173 MVLHPDEIHDGEAGT 187

RESULT 24
US-09-489-039A-12578
; Sequence 12578, Application US/09489039A

```

: Sequence 12578, Application US/05495039A
:
: Patent No. 6610836
:
: GENERAL INFORMATION:
:
: APPLICANT: Gary Bretton et. al
:
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
:
: TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
:
: FILE REFERENCE: 2709.2004001
:
: CURRENT APPLICATION NUMBER: US/09/489,039A
:
: CURRENT FILING DATE: 2000-01-27
:
: PRIOR APPLICATION NUMBER: US 60/117,747
:
: PRIOR FILING DATE: 1999-01-29
:
: NUMBER OF SEQ ID NOS: 14342
:
: SEQ ID NO 12578
:
: LENGTH: 414

```

Query Match	48.7%	Score 37;	DB 2;	Length 414;
Best Local Similarity	53.8%	Pred. No.	1.2e+02;	
Matches	7;	Mismatches	4;	Indels 0;
	7;	Conservative	0;	Gaps 0;

Qy 2 VLRPSRVHNSSEN 14
:: ||| |||
Db , 225 LINPYRVEESEN 237

```

RESULT 25
US-07-816-283-10
; Sequence 10, Application US/07816283
; Patent No. 5436155
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme I.
; APPLICANT: Yamada, Yuichiro
; APPLICANT: Seino, Susumu
; TITLE OF INVENTION: SOMATOSTATIN RECEPTORS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: PO Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/816,283
; FILING DATE: 19911231
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDaniel, C. Steven
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 418 amino acids

```

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-816-283-10

Query Match 48.7%; Score 37; DB 1; Length 418;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSENT 15
DB 327 VLLRPSRRVRSQPT 341

RESULT 26

US-08-417-103-10
Sequence 10, Application US/08417103
Patent No. 5723299
GENERAL INFORMATION:
APPLICANT: Bell, Graeme I.
APPLICANT: Yamada, Yuichiro
APPLICANT: Seino, Suetsumu
TITLE OF INVENTION: SOMATOSTATIN RECEPTORS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,103
FILING DATE: 05-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/816,283
FILING DATE: 01-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: ARCD:144
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-103-10

Query Match 48.7%; Score 37; DB 1; Length 418;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSENT 15
DB 327 VLLRPSRRVRSQPT 341

RESULT 27

US-09-826-509-573
Sequence 573, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Bruinsma, Karin

APPLICANT: Liaw, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6805054-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: PatentIn Version 2.1
SEQ ID NO 573
LENGTH: 418
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-573

Query Match 48.7%; Score 37; DB 2; Length 418;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSENT 15
DB 327 VLLRPSRRVRSQPT 341

RESULT 28

US-08-264-002-7
Sequence 7, Application US/08264002
Patent No. 5559019
GENERAL INFORMATION:
APPLICANT: GUI, JIAN-FANG
APPLICANT: FU, XIANG-DONG
TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 Century Park East, Fifth Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,002
FILING DATE: 22-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TUMARKIN PH.D., LISA A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD3590
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-264-002-7

Query Match 48.7%; Score 37; DB 1; Length 544;
Best Local Similarity 53.3%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 IVLPRSRVHNSSENT 15

Db 316 ISLRDSQKSHSPNS 330
||| | | | | |

RESULT 29

US-09-457-040B-15
; Sequence 15, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Human
US-09-457-040B-15

Query Match 48.7%; Score 37; DB 2; Length 544;
Best Local Similarity 53.3%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEENT 15
||| | | | | |

Db 316 ISLRDSQKSHSPNS 330

RESULT 30

US-09-270-767-45640
; Sequence 45640, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45640
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45640

Query Match 48.7%; Score 37; DB 2; Length 565;
Best Local Similarity 46.2%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSEEN 14
:: | | | | |

Db 9 LMDPTLVHNTREN 21

RESULT 31

US-09-252-991A-27405
; Sequence 27405, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27405
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27405

Query Match 48.7%; Score 37; DB 2; Length 1021;
Best Local Similarity 58.3%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 12
||| | | | | |

Db 407 IVLRPGDAHDAE 418

RESULT 32

US-09-248-796A-24931
; Sequence 24931, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 26208
; SEQ ID NO 24931
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24931

Query Match 47.4%; Score 36; DB 2; Length 62;
Best Local Similarity 40.0%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEENT 15
||| | | | | |

Db 12 ITLTPSKLHKIQTT 26

RESULT 33

US-09-270-767-33410
; Sequence 33410, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33410
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33410

Query Match 47.4%; Score 36; DB 2; Length 172;
Best Local Similarity 58.3%; Pred. No. 68;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSEEN 14
||| | | | | |

Db 121 LRPSRFVYTKEN 132

```

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33440

Query Match      47.4%; Score 36; DB 2; Length 212;
Best Local Similarity 54.5%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      5 PSRVHSEENT 15
      :|||:
Db      42 PARVHPGDHNT 52

RESULT 36
US-09-270-767-48657
; Sequence 48657, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48657
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-48657

Query Match      47.4%; Score 36; DB 2; Length 212;
Best Local Similarity 54.5%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      5 PSRVHSEENT 15
      :|||:
Db      42 PARVHPGDHNT 52

RESULT 37
US-09-252-991A-29516
; Sequence 29516, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29516
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29516

Query Match      47.4%; Score 36; DB 2; Length 219;
Best Local Similarity 54.5%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 IVLRPSRVHNS 11
      :|||:
Db      185 LMLRPGRAHQ 195

RESULT 38
US-09-134-000C-5252

```

```

RESULT 34
US-08-858-207A-381
; Sequence 381, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 381:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-381

Query Match      47.4%; Score 36; DB 2; Length 191;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      2 VLPRPSRVHSEENT 15
      :|||:
Db      122 VLPRGIVHRDKDT 135

RESULT 35
US-09-270-767-33440
; Sequence 33440, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33440
; LENGTH: 212

```

```
; Sequence 5252, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5252
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-5252

Query Match 47.4%; Score 36; DB 2; Length 269;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHSEENT 15
Db 3 LLMKVAIVHSEKS 17
::: : ||||| :

RESULT 39
US-09-190-824-2
; Sequence 2, Application US/09190824A
; Patent No. 6361965
; GENERAL INFORMATION:
; APPLICANT: Powell, David J.
; TITLE OF INVENTION: Yfii pseudouridine synthase
; FILE REFERENCE: GM10181
; CURRENT APPLICATION NUMBER: US/09/190,824A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-190-824-2

Query Match 47.4%; Score 36; DB 2; Length 292;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRPSRVHSEENT 15
Db 122 VLRPGVHRDKDT 135
||||| ||| ::|

RESULT 40
US-09-583-110-5136
; Sequence 5136, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
```

```
; SEQ ID NO 5136
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-5136

Query Match 47.4%; Score 36; DB 2; Length 295;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRPSRVHSEENT 15
Db 122 VLRPGVHRDKDT 135
||||| ||| ::|

RESULT 41
US-09-107-433-3370
; Sequence 3370, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <unknown>
; OPERATING SYSTEM: <unknown>
; SOFTWARE: <unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...295
; SEQUENCE DESCRIPTION: SEQ ID NO: 3370:
US-09-107-433-3370

Query Match 47.4%; Score 36; DB 2; Length 295;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRPSRVHSEENT 15
Db 122 VLRPGVHRDKDT 135
||||| ||| ::|
```

Db 122 VLRPGIVHRIDKDT 135

RESULT 42

US-09-213-053-4

Sequence 4, Application US/09213053

Patent No. 6159477

GENERAL INFORMATION:

APPLICANT: AUDONNET, Jean-Christophe

APPLICANT: BAUDU, Philippe

TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN

TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE

TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS

FILE REFERENCE: 454313-2200

CURRENT APPLICATION NUMBER: US/09/213,053

CURRENT FILING DATE: 1998-12-16

EARLIER APPLICATION NUMBER: 9608242

EARLIER FILING DATE: 1996-06-27

EARLIER APPLICATION NUMBER: PCT/FR97/01115

EARLIER FILING DATE: 1997-06-23

NUMBER OF SEQ ID NOS: 31

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 420

TYPE: PRT

ORGANISM: Canine herpesvirus

US-09-213-053-4

Query Match 47.4%; Score 36; DB 2; Length 420;

Best Local Similarity 50.0%; Pred. No. 1.9e+02;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IIVLRPSRVHSEEN 14

Db 80 ITHPSQIINNEEN 93

RESULT 43

US-08-477-451-23

Sequence 23, Application US/08477451

Patent No. 5928865

GENERAL INFORMATION:

APPLICANT: Covacci, Antonello

TITLE OF INVENTION: Helicobacter Pylori Cagi Region

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,451

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: 0335,002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2708

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 469 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-451-23

Query Match 47.4%; Score 36; DB 1; Length 469;

Best Local Similarity 38.5%; Pred. No. 2.2e+02;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15

Db 49 LKPRKHNEANHT 61

RESULT 44

US-09-934-901-20

Sequence 20, Application US/09934901

Patent No. 6555353

GENERAL INFORMATION:

APPLICANT: Koffas, Mattheos

APPLICANT: Odom, J. Martin

APPLICANT: No. 6555353ton, Kelley C.

APPLICANT: Ye, Rick

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA

CURRENT APPLICATION NUMBER: US/09/934,901

CURRENT FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,906

PRIOR FILING DATE: September 1, 2000

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Microsoft Office 97

SEQ ID NO 20

LENGTH: 501

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-09-934-901-20

Query Match 47.4%; Score 36; DB 2; Length 501;

Best Local Similarity 61.5%; Pred. No. 2.3e+02;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15

Db 283 LRPTVRNVEEKT 295

RESULT 45

US-09-934-868-10

Sequence 10, Application US/09934868

Patent No. 6689601

GENERAL INFORMATION:

APPLICANT: Koffas, Mattheos

APPLICANT: Odom, James M.

APPLICANT: Schenzle, Andreas J.

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1596 US NA

CURRENT APPLICATION NUMBER: US/09/934,868

CURRENT FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,858

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 81

SOFTWARE: Microsoft Office 97

SEQ ID NO 10

LENGTH: 501

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-09-934-868-10

Query Match 47.4%; Score 36; DB 2; Length 501;

Best Local Similarity 61.5%; Pred. No. 2.3e+02;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15

Db 122 VLRPGIVHRIDKDT 135

RESULT 42

US-09-213-053-4

Sequence 4, Application US/09213053

Patent No. 6159477

GENERAL INFORMATION:

APPLICANT: AUDONNET, Jean-Christophe

APPLICANT: BAUDU, Philippe

TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN

TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE

TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS

FILE REFERENCE: 454313-2200

CURRENT APPLICATION NUMBER: US/09/213,053

CURRENT FILING DATE: 1998-12-16

EARLIER APPLICATION NUMBER: 9608242

EARLIER FILING DATE: 1996-06-27

EARLIER APPLICATION NUMBER: PCT/FR97/01115

EARLIER FILING DATE: 1997-06-23

NUMBER OF SEQ ID NOS: 31

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 420

TYPE: PRT

ORGANISM: Canine herpesvirus

US-09-213-053-4

Query Match 47.4%; Score 36; DB 2; Length 420;

Best Local Similarity 50.0%; Pred. No. 1.9e+02;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 IIVLRPSRVHSEEN 14

Db 80 ITHPSQIINNEEN 93

RESULT 43

US-08-477-451-23

Sequence 23, Application US/08477451

Patent No. 5928865

GENERAL INFORMATION:

APPLICANT: Covacci, Antonello

TITLE OF INVENTION: Helicobacter Pylori Cagi Region

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,451

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: 0335,002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2708

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 469 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-451-23

Query Match 47.4%; Score 36; DB 1; Length 469;

Best Local Similarity 38.5%; Pred. No. 2.2e+02;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15

Db 49 LKPRKHNEANHT 61

RESULT 44

US-09-934-901-20

Sequence 20, Application US/09934901

Patent No. 6555353

GENERAL INFORMATION:

APPLICANT: Koffas, Mattheos

APPLICANT: Odom, J. Martin

APPLICANT: No. 6555353ton, Kelley C.

APPLICANT: Ye, Rick

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1619 US NA

CURRENT APPLICATION NUMBER: US/09/934,901

CURRENT FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,906

PRIOR FILING DATE: September 1, 2000

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Microsoft Office 97

SEQ ID NO 20

LENGTH: 501

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-09-934-901-20

Query Match 47.4%; Score 36; DB 2; Length 501;

Best Local Similarity 61.5%; Pred. No. 2.3e+02;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15

Db 283 LRPTVRNVEEKT 295

RESULT 45

US-09-934-868-10

Sequence 10, Application US/09934868

Patent No. 6689601

GENERAL INFORMATION:

APPLICANT: Koffas, Mattheos

APPLICANT: Odom, James M.

APPLICANT: Schenzle, Andreas J.

TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN

FILE REFERENCE: CL1596 US NA

CURRENT APPLICATION NUMBER: US/09/934,868

CURRENT FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/229,858

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 81

SOFTWARE: Microsoft Office 97

SEQ ID NO 10

LENGTH: 501

TYPE: PRT

ORGANISM: METHYLOMONAS SP.

US-09-934-868-10

Query Match 47.4%; Score 36; DB 2; Length 501;

Best Local Similarity 61.5%; Pred. No. 2.3e+02;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRPSRVHSEENT 15


```
Db      283 LRPIVRNVEEKT 295

RESULT 46
US-10-321-210-20
; Sequence 20, Application US/10321210
; Patent No. 6767744
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6767744ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 501
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-20

Query Match      47.4%; Score 36; DB 2; Length 501;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      3 LRPSRVHNSSEENT 15
Db      283 LRPIVRNVEEKT 295

RESULT 47
US-10-320-874-20
; Sequence 20, Application US/10320874
; Patent No. 6773905
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6773905ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,874
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 501
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-320-874-20

Query Match      47.4%; Score 36; DB 2; Length 501;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      3 LRPSRVHNSSEENT 15
Db      283 LRPIVRNVEEKT 295

RESULT 48
US-10-320-924-20
; Sequence 20, Application US/10320924
; Patent No. 6908992
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6908992ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,924
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 501
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-320-924-20

Query Match      47.4%; Score 36; DB 2; Length 501;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      3 LRPSRVHNSSEENT 15
Db      283 LRPIVRNVEEKT 295

RESULT 49
US-10-701-200-10
; Sequence 10, Application US/10701200
; Patent No. 6958222
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/10/701,200
; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 501
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-701-200-10

Query Match      47.4%; Score 36; DB 2; Length 501;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      3 LRPSRVHNSSEENT 15
Db      283 LRPIVRNVEEKT 295

RESULT 50
US-08-313-553-15
; Sequence 15, Application US/08313553
; Patent No. 5641650
; GENERAL INFORMATION:
; APPLICANT: TURNER, George J.
; APPLICANT: BETLACH, Mary C.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
```

;; TITLE OF INVENTION: IN HALOBACTERIA
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Walter H. Dreger
;; STREET: 4 Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/313,553
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION NUMBER:
;; APPLICATION NUMBER: US/08/038,662
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dreger, Walter H.
;; REGISTRATION NUMBER: 24,190
;; REFERENCE/DOCKET NUMBER: A-57669/WHd
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 398-3249
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 559 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-313-553-15

Query Match 47.4%; Score 36; DB 1; Length 559;
Best Local Similarity 53.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
||| | :||| |
Db 496 VLRASDLHNKAKN 508

RESULT 51
US-08-767-993-15
; Sequence 15, Application US/08767993
; Patent No. 6010885
; GENERAL INFORMATION:
; APPLICANT: TURNER, George J.
; APPLICANT: BETLACH, Mary C.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
; TITLE OF INVENTION: IN HALOBACTERIA
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,993
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.

;; REGISTRATION NUMBER: 24,190
;; REFERENCE/DOCKET NUMBER: A-57669/WHd
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 781-1989
;; TELEFAX: (415) 398-3249
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 559 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-767-993-15

Query Match 47.4%; Score 36; DB 2; Length 559;
Best Local Similarity 53.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 14
||| | :||| |
Db 496 VLRASDLHNKAKN 508

RESULT 52
US-09-540-236-2083
; Sequence 2083, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2083
; LENGTH: 645
; TYPE: PRT
; ORGANISM: M.catarrhalis
; US-09-540-236-2083

Query Match 47.4%; Score 36; DB 2; Length 645;
Best Local Similarity 35.7%; Pred. No. 3.1e+02;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSEN 15
:::|:|:|:|
Db 624 MIQPANITNHDENT 637

RESULT 53
US-09-248-796A-20015
; Sequence 20015, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20015
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-20015

Query Match 47.4%; Score 36; DB 2; Length 653;
Best Local Similarity 54.5%; Pred. No. 3.2e+02;

```
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 4 RPSRVHNSSEN 14
:|||||:
Db 500 KPSRGHKDQEN 510
:|||||:

RESULT 54
US-09-252-991A-20334
; Sequence 20334, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20334
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20334

Query Match 47.4%; Score 36; DB 2; Length 1089;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 4; Gaps 1;
Qy 2 VLRPSRVHNSSEN 15
:|||||:
Db 58 VLRPSRIRQAACSGE 75
:|||||:

RESULT 55
US-08-441-139-9
; Sequence 9, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Wittrup, Dr. Karl D.
; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441.139
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
```

```
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-139-9

Query Match 47.4%; Score 36; DB 1; Length 1250;
Best Local Similarity 53.8%; Pred. No. 6.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 VLRPSRVHNSSEN 14
:|||||:
Db 561 LLRPSQOANSNN 573
:|||||:

RESULT 56
US-09-487-558B-364
; Sequence 364, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 364
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-364

Query Match 47.4%; Score 36; DB 2; Length 1250;
Best Local Similarity 53.8%; Pred. No. 6.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 VLRPSRVHNSSEN 14
:|||||:
Db 561 LLRPSQOANSNN 573
:|||||:

RESULT 57
US-09-213-053-2
; Sequence 2, Application US/09213053
; Patent No. 6159477
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BAUDU, Philippe
; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
; TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200
; CURRENT APPLICATION NUMBER: US/09/213,053
; CURRENT FILING DATE: 1998-12-16
```

EARLIER APPLICATION NUMBER: 9608242
EARLIER FILING DATE: 1996-06-27
EARLIER APPLICATION NUMBER: PCT/FR97/01115
EARLIER FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 2
LENGTH: 1394
TYPE: PRT
ORGANISM: Canine herpesvirus
US-09-213-053-2

Query Match 47.4%; Score 36; DB 2; Length 1394;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSEN 14
| | | | | : | | | |
Db 80 ITNHPISQIINNEEN 93

RESULT 58
US-09-252-991A-17657
Sequence 17657, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17657
LENGTH: 1467
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17657

Query Match 46.7%; Score 35.5; DB 2; Length 1467;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 3 LRPSRVHNSSEN 15
| | | | | : | | | |
Db 23 LRPSRVHNSSEN 34

RESULT 59
US-09-513-999C-6876
Sequence 6876, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513.999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 6876
LENGTH: 72
TYPE: PRT
ORGANISM: Homo sapiens

US-09-513-999C-6876

Query Match 46.1%; Score 35; DB 2; Length 72;
Best Local Similarity 58.3%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSEN 14
| | | | | : | | | |
Db 40 LRPSRVHNSSEN 51

RESULT 60
US-09-471-276-1285
Sequence 1285, Application US/09471276
Patent No. 6822072
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: GENSET.025CP1
CURRENT APPLICATION NUMBER: US/09/471.276
CURRENT FILING DATE: 1999-12-21
EARLIER APPLICATION NUMBER: 09/057,719
EARLIER FILING DATE: 1998-04-09
EARLIER APPLICATION NUMBER: 09/069,047
EARLIER FILING DATE: 1998-04-28
EARLIER APPLICATION NUMBER: PCT/IB99/00712
EARLIER FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 1622
SOFTWARE: Patent.pm
SEQ ID NO 1285
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -18..-1
US-09-471-276-1285

Query Match 46.1%; Score 35; DB 2; Length 73;
Best Local Similarity 58.3%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSEN 14
| | | | | : | | | |
Db 40 LRPSRVHNSSEN 51

Search completed: March 11, 2006, 12:27:12
Job time : 39.5526 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:26:08 ; Search time 117.632 Seconds
(without alignments)
53.280 Million cell updates/sec

Title: US-10-774-242A-1

Perfect score: 76

Sequence: 1 IVLPRSRVHNEENT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	15	4	US-10-774-242-1
2	76	100.0	760	3	US-09-265-606-2
3	76	100.0	760	4	US-10-177-293-136
4	76	100.0	760	4	US-10-301-822-55
5	76	100.0	760	5	US-10-723-860-4171
6	76	100.0	760	5	US-10-884-070A-13
7	48	63.2	120	4	US-10-425-115-320178
8	42	55.3	425	4	US-10-435-696-95
9	41	53.9	108	4	US-10-425-115-361324
10	41	53.9	113	4	US-10-425-115-333380
11	41	53.9	258	4	US-10-424-599-239911
12	41	53.9	458	4	US-10-425-114-64668
13	41	53.9	499	4	US-10-282-122A-48505
14	41	53.9	2828	3	US-09-802-318-21
15	41	53.9	2828	3	US-09-905-129-21
16	41	53.9	2828	3	US-09-991-630-21
17	41	53.9	2828	4	US-10-176-847-54
18	41	53.9	2828	4	US-10-177-293-110
19	41	53.9	2828	4	US-10-301-822-49
20	41	53.9	2828	4	US-10-032-189-126
21	41	53.9	2828	4	US-10-295-027-58
22	41	53.9	2828	4	US-10-295-027-1175
23	41	53.9	2828	4	US-10-072-012-607
24	41	53.9	2828	4	US-10-072-012-608
25	41	53.9	2828	4	US-10-454-351-21
26	41	53.9	2828	5	US-10-723-860-554
27	41	53.9	2993	5	US-10-450-763-54243

5266	5	US-10-450-763-38920	Sequence 38920, A
81	4	US-10-767-701-62671	Sequence 62671, A
90	4	US-10-314-669-133	Sequence 133, App
90	4	US-10-669-861-131	Sequence 133, App
90	5	US-10-732-620-119	Sequence 119, App
115	3	US-09-738-626-5581	Sequence 5581, App
210	4	US-10-314-669-221	Sequence 221, App
210	4	US-10-669-861-221	Sequence 221, App
238	4	US-10-314-669-222	Sequence 222, App
238	4	US-10-314-669-223	Sequence 223, App
238	4	US-10-314-669-224	Sequence 224, App
238	4	US-10-669-861-222	Sequence 223, App
238	4	US-10-669-861-223	Sequence 224, App
238	4	US-10-669-861-224	Sequence 224, App
313	4	US-10-425-114-72739	Sequence 72739, A
313	4	US-10-437-963-157183	Sequence 157183, A
836	4	US-10-149-310-106	Sequence 106, App
975	6	US-11-097-143-40320	Sequence 40320, A
13	5	US-10-898-141-17	Sequence 17, Appl
124	4	US-10-425-115-253664	Sequence 253664, A
417	4	US-10-437-963-103361	Sequence 103361, A
421	5	US-10-450-763-58555	Sequence 58555, A
497	3	US-09-796-753-116	Sequence 116, App
946	4	US-10-437-963-172291	Sequence 172291, A
60	4	US-10-425-115-296047	Sequence 296047, A
122	4	US-10-424-599-232656	Sequence 232656, A
148	4	US-10-108-260A-2889	Sequence 2889, App
183	3	US-09-864-761-35808	Sequence 35808, A
232	3	US-09-815-242-5560	Sequence 5560, App
232	3	US-09-815-242-12569	Sequence 12569, A
235	4	US-10-767-701-32297	Sequence 32297, A
242	4	US-10-437-963-204794	Sequence 204794, A
303	4	US-10-767-701-35712	Sequence 35712, A
303	4	US-10-425-115-345754	Sequence 345754, A
308	5	US-10-450-763-33746	Sequence 33746, A
312	4	US-10-425-114-61210	Sequence 61210, A
347	4	US-10-425-114-82553	Sequence 82553, A
368	4	US-10-369-493-21985	Sequence 21985, A
371	4	US-10-076-157-6	Sequence 6, Appl
452	4	US-10-425-115-229873	Sequence 229873, A
495	6	US-11-097-143-6078	Sequence 6078, App
798	5	US-10-450-763-49865	Sequence 49865, A
948	5	US-10-732-923-23243	Sequence 23243, A
966	4	US-10-369-493-2172	Sequence 2172, App
2652	4	US-10-454-351-30	Sequence 30, Appl
2845	4	US-10-093-463-2	Sequence 2, Appl
317	4	US-10-437-963-158614	Sequence 158614, A
74	6	US-11-097-143-31629	Sequence 31629, A
150	4	US-10-424-599-242516	Sequence 242516, A
155	4	US-10-437-963-153739	Sequence 153739, A
163	4	US-10-767-701-37860	Sequence 37860, A
185	4	US-10-437-963-173911	Sequence 173911, A
219	5	US-10-739-930-9541	Sequence 9541, App
301	4	US-10-369-493-6458	Sequence 6458, App
304	4	US-10-437-963-129502	Sequence 129502, A
304	4	US-10-072-012-727	Sequence 727, App
345	4	US-10-153-739-2	Sequence 2, Appl
345	4	US-10-264-237-2583	Sequence 2583, App
366	5	US-10-909-436-4	Sequence 4, Appl
405	4	US-10-283-122A-48496	Sequence 48496, A
408	4	US-10-377-636-2	Sequence 2, Appl
418	3	US-09-992-331-17	Sequence 17, Appl
418	3	US-09-943-007A-2	Sequence 2, Appl
418	3	US-09-964-923A-18	Sequence 18, Appl
418	3	US-09-826-509-573	Sequence 573, App
418	4	US-10-225-567A-316	Sequence 316, App
418	4	US-10-262-313-17	Sequence 17, Appl
418	4	US-10-680-402-18	Sequence 18, Appl
418	4	US-10-768-878-17	Sequence 17, Appl
418	4	US-10-923-095-573	Sequence 573, App
455	4	US-10-425-115-196898	Sequence 196898, A
458	4	US-10-369-493-3910	Sequence 3910, App
544	4	US-10-369-493-2376	Sequence 2376, App

101	37	48.7	737.4	US-10-282-122A-78534	Sequence 78534, A	174	35	46.1	52	4	US-10-437-963-126608	Sequence 126608,
102	37	48.7	737.5	US-10-732-923-20096	Sequence 20096, A	175	35	46.1	55	3	US-09-864-408A-20	Sequence 20, Appl
103	37	48.7	751.5	US-10-732-923-20094	Sequence 20094, A	176	35	46.1	66	4	US-10-425-115-262642	Sequence 262642,
104	37	48.7	805.5	US-10-739-930-6331	Sequence 6331, A	177	35	46.1	67	5	US-10-450-763-39949	Sequence 39949, A
105	37	48.7	1006.4	US-10-437-963-155082	Sequence 155082, A	178	35	46.1	72	4	US-10-424-599-241134	Sequence 241134,
106	37	48.7	1356.4	US-11-097-143-13869	Sequence 13869, A	179	35	46.1	73	4	US-10-425-115-231596	Sequence 231596,
107	37	48.7	4498.4	US-10-712-124-68	Sequence 68, Appl	180	35	46.1	73	5	US-10-926-683-1285	Sequence 1285, A
108	37	48.7	4498.6	US-11-097-143-2577	Sequence 2577, A	181	35	46.1	73	5	US-10-450-763-48189	Sequence 48189, A
109	36.5	48.0	87.4	US-10-425-115-293327	Sequence 293327, A	182	35	46.1	77	4	US-10-424-599-185699	Sequence 185699,
110	36.5	48.0	355.4	US-10-363-829-390	Sequence 390, App	183	35	46.1	83	4	US-10-425-115-283163	Sequence 283163,
111	36	47.4	65.4	US-10-437-963-202462	Sequence 202462, A	184	35	46.1	83	4	US-10-425-115-286438	Sequence 286438,
112	36	47.4	77.4	US-10-424-599-170323	Sequence 170323, A	185	35	46.1	84	4	US-10-437-963-127934	Sequence 127934,
113	36	47.4	79.4	US-10-424-599-167881	Sequence 167881, A	186	35	46.1	88	3	US-09-833-245-1900	Sequence 1900, A
114	36	47.4	100.4	US-10-425-115-213439	Sequence 213439, A	187	35	46.1	88	3	US-09-833-245-1901	Sequence 1901, A
115	36	47.4	105.4	US-10-425-115-338472	Sequence 338472, A	188	35	46.1	94	4	US-10-767-701-33937	Sequence 33937, A
116	36	47.4	119.4	US-10-424-599-226433	Sequence 226433, A	189	35	46.1	100	4	US-10-425-115-368544	Sequence 368544, A
117	36	47.4	146.4	US-10-424-599-237494	Sequence 237494, A	190	35	46.1	108	5	US-10-450-763-32028	Sequence 32028, A
118	36	47.4	173.4	US-10-369-493-11804	Sequence 11804, A	191	35	46.1	116	4	US-10-425-115-367724	Sequence 367724, A
119	36	47.4	176.4	US-10-425-115-222811	Sequence 222811, A	192	35	46.1	117	3	US-09-833-245-1898	Sequence 1898, A
120	36	47.4	205.4	US-10-464-368-49	Sequence 49, Appl	193	35	46.1	126	4	US-10-425-115-366330	Sequence 366330, A
121	36	47.4	211.4	US-10-029-386-34052	Sequence 34052, A	194	35	46.1	138	4	US-10-767-701-39856	Sequence 39856, A
122	36	47.4	221.5	US-10-450-763-59380	Sequence 59380, A	195	35	46.1	138	5	US-10-450-763-35463	Sequence 35463, A
123	36	47.4	242.3	US-09-815-242-13051	Sequence 13051, A	196	35	46.1	138	5	US-10-450-763-44781	Sequence 44781, A
124	36	47.4	251.3	US-09-815-242-12717	Sequence 12717, A	197	35	46.1	138	5	US-10-450-763-45098	Sequence 45098, A
125	36	47.4	295.3	US-09-815-242-13351	Sequence 13351, A	198	35	46.1	151	4	US-10-425-115-216221	Sequence 216221, A
126	36	47.4	295.3	US-09-815-242-13691	Sequence 13691, A	199	35	46.1	152	4	US-10-424-599-276302	Sequence 276302, A
127	36	47.4	295.5	US-10-472-928-1814	Sequence 1814, A	200	35	46.1	160	4	US-10-424-599-15703	Sequence 15703, A
128	36	47.4	398.4	US-10-617-320-3370	Sequence 3370, A	201	35	46.1	170	4	US-10-425-115-296693	Sequence 296693, A
129	36	47.4	407.4	US-10-369-493-23676	Sequence 23676, A	202	35	46.1	178	4	US-10-425-115-212520	Sequence 212520, A
130	36	47.4	332.4	US-10-424-599-216012	Sequence 216012, A	203	35	46.1	182	4	US-10-767-701-62722	Sequence 62722, A
131	36	47.4	365.4	US-10-437-963-148945	Sequence 148945, A	204	35	46.1	192	4	US-10-437-963-177175	Sequence 177175, A
132	36	47.4	387.4	US-10-424-599-267453	Sequence 267453, A	205	35	46.1	197	4	US-10-424-599-207929	Sequence 207929, A
133	36	47.4	398.4	US-10-156-761-13711	Sequence 13711, A	206	35	46.1	202	5	US-10-732-923-15827	Sequence 15827, A
134	36	47.4	407.4	US-10-369-493-3506	Sequence 3506, A	207	35	46.1	206	5	US-10-450-763-38582	Sequence 38582, A
135	36	47.4	423.3	US-09-815-242-5873	Sequence 5873, A	208	35	46.1	210	4	US-10-299-636-46	Sequence 27963, A
136	36	47.4	436.4	US-10-038-854-130	Sequence 130, App	209	35	46.1	213	6	US-11-097-143-27963	Sequence 145610, A
137	36	47.4	441.4	US-10-094-749-2984	Sequence 2984, A	210	35	46.1	223	4	US-10-437-963-145610	Sequence 145610, A
138	36	47.4	441.4	US-10-479-435-25	Sequence 25, Appl	211	35	46.1	225	4	US-10-424-599-165239	Sequence 165239, A
139	36	47.4	460.4	US-10-437-963-202505	Sequence 202505, A	212	35	46.1	232	4	US-10-282-122A-44063	Sequence 44063, A
140	36	47.4	491.4	US-10-437-963-147211	Sequence 147211, A	213	35	46.1	237	3	US-09-864-761-33860	Sequence 33860, A
141	36	47.4	501.3	US-09-934-901-20	Sequence 20, Appl	214	35	46.1	241	4	US-10-437-963-104406	Sequence 104406, A
142	36	47.4	501.3	US-09-934-868-10	Sequence 10, Appl	215	35	46.1	246	5	US-10-732-923-14079	Sequence 14079, A
143	36	47.4	501.4	US-10-320-924-20	Sequence 20, Appl	216	35	46.1	270	4	US-10-094-749-2962	Sequence 2962, A
144	36	47.4	501.4	US-10-320-874-20	Sequence 20, Appl	217	35	46.1	283	4	US-10-424-599-173349	Sequence 173349, A
145	36	47.4	501.4	US-10-321-210-20	Sequence 20, Appl	218	35	46.1	298	4	US-10-424-599-165240	Sequence 165240, A
146	36	47.4	501.4	US-10-701-200-10	Sequence 10, Appl	219	35	46.1	314	4	US-10-369-493-22182	Sequence 22182, A
147	36	47.4	501.4	US-10-437-963-177928	Sequence 177928, A	220	35	46.1	316	4	US-10-425-115-221659	Sequence 221659, A
148	36	47.4	527.4	US-10-424-599-252225	Sequence 252225, A	221	35	46.1	321	4	US-10-425-114-61280	Sequence 61280, A
149	36	47.4	547.4	US-10-424-599-252225	Sequence 252225, A	222	35	46.1	330	4	US-10-767-701-42700	Sequence 42700, A
150	36	47.4	551.4	US-10-282-122A-52951	Sequence 52951, A	223	35	46.1	336	3	US-09-780-669-940	Sequence 940, App
151	36	47.4	552.4	US-10-282-122A-51769	Sequence 51769, A	224	35	46.1	336	3	US-09-822-827-940	Sequence 940, App
152	36	47.4	579.4	US-10-282-122A-44483	Sequence 44483, A	225	35	46.1	336	3	US-09-895-793-940	Sequence 940, App
153	36	47.4	592.4	US-10-450-763-54636	Sequence 54636, A	226	35	46.1	336	3	US-09-895-814-940	Sequence 940, App
154	36	47.4	606.4	US-10-094-749-3064	Sequence 3064, A	227	35	46.1	336	4	US-10-012-896-940	Sequence 940, App
155	36	47.4	613.4	US-10-450-763-56313	Sequence 56313, A	228	35	46.1	336	4	US-10-144-678A-940	Sequence 940, App
156	36	47.4	627.4	US-11-097-143-11217	Sequence 11217, A	229	35	46.1	336	4	US-10-294-025-940	Sequence 940, App
157	36	47.4	691.4	US-10-437-963-136688	Sequence 136688, A	230	35	46.1	347	3	US-09-759-143-590	Sequence 590, App
158	36	47.4	1066.4	US-10-369-493-22668	Sequence 22668, A	231	35	46.1	347	3	US-09-822-827-590	Sequence 590, App
159	36	47.4	1250.4	US-09-801-368-364	Sequence 364, App	232	35	46.1	347	3	US-09-895-793-590	Sequence 590, App
160	36	47.4	1250.4	US-10-369-493-1672	Sequence 1672, A	233	35	46.1	347	3	US-09-895-814-590	Sequence 590, App
161	36	47.4	2162.4	US-10-267-502-296	Sequence 296, App	234	35	46.1	347	4	US-10-012-896-590	Sequence 590, App
162	36	47.4	2162.6	US-11-097-143-20514	Sequence 20514, A	235	35	46.1	347	4	US-10-144-678A-590	Sequence 590, App
163	35.5	46.7	683.3	US-09-792-630-39	Sequence 39, Appl	236	35	46.1	347	4	US-10-294-025-590	Sequence 590, App
164	35.5	46.7	683.3	US-09-953-351-39	Sequence 39, Appl	237	35	46.1	347	4	US-10-424-599-279885	Sequence 279885, A
165	35.5	46.7	683.4	US-10-080-376-39	Sequence 39, Appl	238	35	46.1	359	4	US-10-424-599-279885	Sequence 18, Appl
166	35.5	46.7	683.4	US-10-082-671-45	Sequence 45, Appl	239	35	46.1	364	5	US-10-737-318-18	Sequence 72366, A
167	35.5	46.7	683.4	US-10-097-100-39	Sequence 39, Appl	240	35	46.1	374	3	US-09-976-165-25	Sequence 25, Appl
168	35.5	46.7	683.4	US-10-023-208-39	Sequence 39, Appl	241	35	46.1	374	3	US-09-227-853A-2	Sequence 9, Appl
169	35.5	46.7	13.5	US-10-898-141-8	Sequence 8, Appl	242	35	46.1	374	4	US-10-052-586-406	Sequence 406, App
170	35	46.1	42.4	US-10-424-599-215047	Sequence 215047, A	243	35	46.1	374	4	US-10-174-590-406	Sequence 406, App
171	35	46.1	47.4	US-10-424-599-205739	Sequence 205739, A	244	35	46.1	374	4	US-10-176-758-406	Sequence 406, App
172	35	46.1	48.4	US-10-437-963-119308	Sequence 119308, A	245	35	46.1	374	4		
173	35	46.1	48.4	US-10-425-115-361654	Sequence 361654, A	246	35	46.1	374	4		

393	35	46.1	374	4	US-10-188-775-406	Sequence 406, App	466	35	46.1	374	4	US-10-187-752-406	Sequence 406, App
394	35	46.1	374	4	US-10-194-462-406	Sequence 406, App	467	35	46.1	374	4	US-10-187-887-406	Sequence 406, App
395	35	46.1	374	4	US-10-196-745-406	Sequence 406, App	468	35	46.1	374	4	US-10-194-461-406	Sequence 406, App
396	35	46.1	374	4	US-10-196-762-406	Sequence 406, App	469	35	46.1	374	4	US-10-195-892-406	Sequence 406, App
397	35	46.1	374	4	US-10-197-695-406	Sequence 406, App	470	35	46.1	374	4	US-10-196-751-406	Sequence 406, App
398	35	46.1	374	4	US-10-195-894-406	Sequence 406, App	471	35	46.1	374	4	US-10-197-694-406	Sequence 406, App
399	35	46.1	374	4	US-10-176-484-406	Sequence 406, App	472	35	46.1	374	4	US-10-197-697-406	Sequence 406, App
400	35	46.1	374	4	US-10-176-753-406	Sequence 406, App	473	35	46.1	374	4	US-10-197-707-406	Sequence 406, App
401	35	46.1	374	4	US-10-176-317-406	Sequence 406, App	474	35	46.1	374	4	US-10-199-303-406	Sequence 406, App
402	35	46.1	374	4	US-10-176-982-406	Sequence 406, App	475	35	46.1	374	4	US-10-199-318-406	Sequence 406, App
403	35	46.1	374	4	US-10-179-506-406	Sequence 406, App	476	35	46.1	374	4	US-10-199-458-406	Sequence 406, App
404	35	46.1	374	4	US-10-179-513-406	Sequence 406, App	477	35	46.1	374	4	US-10-199-462-406	Sequence 406, App
405	35	46.1	374	4	US-10-179-514-406	Sequence 406, App	478	35	46.1	374	4	US-10-201-324-406	Sequence 406, App
406	35	46.1	374	4	US-10-179-522-406	Sequence 406, App	479	35	46.1	374	4	US-10-201-328-406	Sequence 406, App
407	35	46.1	374	4	US-10-180-556-406	Sequence 406, App	480	35	46.1	374	4	US-10-201-527-406	Sequence 406, App
408	35	46.1	374	4	US-10-180-560-406	Sequence 406, App	481	35	46.1	374	4	US-10-201-528-406	Sequence 406, App
409	35	46.1	374	4	US-10-183-015-406	Sequence 406, App	482	35	46.1	374	4	US-10-201-529-406	Sequence 406, App
410	35	46.1	374	4	US-10-184-615-406	Sequence 406, App	483	35	46.1	374	4	US-10-201-530-406	Sequence 406, App
411	35	46.1	374	4	US-10-184-620-406	Sequence 406, App	484	35	46.1	374	4	US-10-202-408-406	Sequence 406, App
412	35	46.1	374	4	US-10-184-643-406	Sequence 406, App	485	35	46.1	374	4	US-10-202-409-406	Sequence 406, App
413	35	46.1	374	4	US-10-184-656-406	Sequence 406, App	486	35	46.1	374	4	US-10-202-411-406	Sequence 406, App
414	35	46.1	374	4	US-10-192-010-406	Sequence 406, App	487	35	46.1	374	4	US-10-202-472-406	Sequence 406, App
415	35	46.1	374	4	US-10-205-908-406	Sequence 406, App	488	35	46.1	374	4	US-10-205-502-406	Sequence 406, App
416	35	46.1	374	4	US-10-184-619-406	Sequence 406, App	489	35	46.1	374	4	US-10-205-507-406	Sequence 406, App
417	35	46.1	374	4	US-10-187-599-406	Sequence 406, App	490	35	46.1	374	4	US-10-205-511-406	Sequence 406, App
418	35	46.1	374	4	US-10-187-750-406	Sequence 406, App	491	35	46.1	374	4	US-10-205-902-406	Sequence 406, App
419	35	46.1	374	4	US-10-188-780-406	Sequence 406, App	492	35	46.1	374	4	US-10-205-907-406	Sequence 406, App
420	35	46.1	374	4	US-10-192-015-406	Sequence 406, App	493	35	46.1	374	4	US-10-176-484-406	Sequence 406, App
421	35	46.1	374	4	US-10-194-394-406	Sequence 406, App	494	35	46.1	374	4	US-10-194-456-406	Sequence 406, App
422	35	46.1	374	4	US-10-194-425-406	Sequence 406, App	495	35	46.1	374	4	US-10-196-758-406	Sequence 406, App
423	35	46.1	374	4	US-10-194-485-406	Sequence 406, App	496	35	46.1	374	4	US-10-198-770-406	Sequence 406, App
424	35	46.1	374	4	US-10-195-885-406	Sequence 406, App	497	35	46.1	374	4	US-10-199-308-406	Sequence 406, App
425	35	46.1	374	4	US-10-195-899-406	Sequence 406, App	498	35	46.1	374	4	US-10-200-617-406	Sequence 406, App
426	35	46.1	374	4	US-10-196-748-406	Sequence 406, App	499	35	46.1	374	4	US-10-205-893-406	Sequence 406, App
427	35	46.1	374	4	US-10-196-750-406	Sequence 406, App	500	35	46.1	374	4	US-10-205-897-406	Sequence 406, App
428	35	46.1	374	4	US-10-197-699-406	Sequence 406, App							
429	35	46.1	374	4	US-10-197-700-406	Sequence 406, App							
430	35	46.1	374	4	US-10-197-705-406	Sequence 406, App							
431	35	46.1	374	4	US-10-198-764-406	Sequence 406, App							
432	35	46.1	374	4	US-10-198-765-406	Sequence 406, App							
433	35	46.1	374	4	US-10-198-768-406	Sequence 406, App							
434	35	46.1	374	4	US-10-198-769-406	Sequence 406, App							
435	35	46.1	374	4	US-10-198-795-406	Sequence 406, App							
436	35	46.1	374	4	US-10-199-305-406	Sequence 406, App							
437	35	46.1	374	4	US-10-199-306-406	Sequence 406, App							
438	35	46.1	374	4	US-10-199-310-406	Sequence 406, App							
439	35	46.1	374	4	US-10-199-311-406	Sequence 406, App							
440	35	46.1	374	4	US-10-199-314-406	Sequence 406, App							
441	35	46.1	374	4	US-10-199-317-406	Sequence 406, App							
442	35	46.1	374	4	US-10-199-665-406	Sequence 406, App							
443	35	46.1	374	4	US-10-199-666-406	Sequence 406, App							
444	35	46.1	374	4	US-10-199-669-406	Sequence 406, App							
445	35	46.1	374	4	US-10-201-534-406	Sequence 406, App							
446	35	46.1	374	4	US-10-201-770-406	Sequence 406, App							
447	35	46.1	374	4	US-10-201-855-406	Sequence 406, App							
448	35	46.1	374	4	US-10-201-856-406	Sequence 406, App							
449	35	46.1	374	4	US-10-202-469-406	Sequence 406, App							
450	35	46.1	374	4	US-10-202-470-406	Sequence 406, App							
451	35	46.1	374	4	US-10-202-476-406	Sequence 406, App							
452	35	46.1	374	4	US-10-202-934-406	Sequence 406, App							
453	35	46.1	374	4	US-10-202-935-406	Sequence 406, App							
454	35	46.1	374	4	US-10-202-936-406	Sequence 406, App							
455	35	46.1	374	4	US-10-202-939-406	Sequence 406, App							
456	35	46.1	374	4	US-10-205-504-406	Sequence 406, App							
457	35	46.1	374	4	US-10-205-509-406	Sequence 406, App							
458	35	46.1	374	4	US-10-205-895-406	Sequence 406, App							
459	35	46.1	374	4	US-10-205-899-406	Sequence 406, App							
460	35	46.1	374	4	US-10-205-900-406	Sequence 406, App							
461	35	46.1	374	4	US-10-205-909-406	Sequence 406, App							
462	35	46.1	374	4	US-10-195-890-406	Sequence 406, App							
463	35	46.1	374	4	US-10-183-002-406	Sequence 406, App							
464	35	46.1	374	4	US-10-184-621-406	Sequence 406, App							
465	35	46.1	374	4	US-10-184-638-406	Sequence 406, App							

ALIGNMENTS

RESULT 1

US-10-774-242-1

Sequence 1, Application US/10774242

Publication No. US20040203102A1

GENERAL INFORMATION:

APPLICANT: McKee, Patrick A

APPLICANT: Lee, Kyung N.

APPLICANT: Jackson, Kenneth W.

APPLICANT: Christiansen, Victoria J.

TITLE OF INVENTION: ANTIPLASMIN CLEAVING ENZYME

FILE REFERENCE: 5820.646

CURRENT APPLICATION NUMBER: US/10/774,242

CURRENT FILING DATE: 2004-02-06

PRIOR APPLICATION NUMBER: 60/445,774

PRIOR FILING DATE: 2003-02-07

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 15

TYPE: PRT

ORGANISM: Homo sapiens

US-10-774-242-1

Query Match

Best Local Similarity

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Score 76; DB 4; Length 15;

Pred. No. 3.4e-06;

Qy

1

IVLRPSRVHSEENT 15

|||||

Db

1

IVLRPSRVHSEENT 15

|||||

RESULT 2

```
US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789Alman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-2

Query Match 100.0%; Score 76; DB 3; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEENT 15
DB 24 IVLRPSRVHNSSEENT 38

RESULT 3
US-10-177-293-136
; Sequence 136, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: East Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
```

```
; APPLICANT: Mexic, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-136

Query Match 100.0%; Score 76; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVLRPSRVHNSSEENT 15
DB 24 IVLRPSRVHNSSEENT 38

RESULT 4
US-10-301-822-55
; Sequence 55, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-55

Query Match 100.0%; Score 76; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 361324
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_92705C.1.pap
US-10-425-115-361324

Query Match 53.9%; Score 41; DB 4; Length 108;
Best Local Similarity 45.0%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 2; Indels 6; Gaps 1;

Qy 2 VLKPS-----RVHNSSEET 15
||||| :|||:|
Db 36 LLKPSAIPTHGRHNTERTT 55

RESULT 10
US-10-425-115-353380
; Sequence 353380, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 353380
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(113)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_85457C.1.pap
US-10-425-115-353380

Query Match 53.9%; Score 41; DB 4; Length 113;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNS 11
||||| :|||:|
Db 88 IILRPGSRSHN 98

RESULT 11
US-10-424-599-239911
; Sequence 239911, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 239911
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(258)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_58665C.1.pap
US-10-424-599-239911

Query Match 53.9%; Score 41; DB 4; Length 258;
Best Local Similarity 57.1%; Pred. No. 94;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSEEN 14
||||| :|||:|
Db 103 IVLMPETVHRSEQS 116

RESULT 12
US-10-425-114-64668
; Sequence 64668, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64668
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4571-006-D12_FLI.pap
US-10-425-114-64668

Query Match 53.9%; Score 41; DB 4; Length 458;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHN 10
||||| :|||:|
Db 332 IVLLPSRMHN 341

RESULT 13
US-10-282-122A-48505
; Sequence 48505, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

Matches	5;	Conservative	7;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	IVLRPSRVHNSEEN 14							
Db	2238	VNMKPAKIEHKEEN 2251							
RESULT 15									
US-09-905-129-21									
; Sequence 21, Application US/09905129									
; Patent No. US20020137705A1									
; GENERAL INFORMATION:									
; APPLICANT: Einat, et al									
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODU									
; FILE REFERENCE: 540579-2007.2									
; CURRENT APPLICATION NUMBER: US/09/905,129									
; CURRENT FILING DATE: 2001-07-13									
; PRIOR APPLICATION NUMBER: 09/802,318									
; PRIOR FILING DATE: 2001-03-08									
; PRIOR APPLICATION NUMBER: 60/207,821									
; PRIOR FILING DATE: 2000-05-30									
; PRIOR APPLICATION NUMBER: 60/084,944									
; PRIOR FILING DATE: 1998-05-11									
; PRIOR APPLICATION NUMBER: 60/085,673									
; PRIOR FILING DATE: 1998-05-15									
; NUMBER OF SEQ ID NOS: 25									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 21									
; LENGTH: 2828									
; TYPE: PRT									
; ORGANISM: homo sapiens									
US-09-905-129-21									
Query Match 53.9%; Score 41; DB 3; Length 2828;									
Best Local Similarity 35.7%; Pred. No. 1.3e+03;									
Matches	5;	Conservative	7;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	IVLRPSRVHNSEEN 14							
Db	2238	VNMKPAKIEHKEEN 2251							
RESULT 16									
US-09-991-630-21									
; Sequence 21, Application US/09991630									
; Patent No. US20020151514A1									
; GENERAL INFORMATION:									
; APPLICANT: Einat, et al									
; TITLE OF INVENTION: GENES ASSOCIATED WITH MECHANICAL STRESS, EXPRESSION PRODU									
; FILE REFERENCE: 540579-2007.3									
; CURRENT APPLICATION NUMBER: US/09/991,630									
; CURRENT FILING DATE: 2001-11-06									
; PRIOR APPLICATION NUMBER: 09/905,129									
; PRIOR FILING DATE: 2001-07-13									
; PRIOR APPLICATION NUMBER: 09/802,318									
; PRIOR FILING DATE: 2001-03-08									
; PRIOR APPLICATION NUMBER: 09/729,485									
; PRIOR FILING DATE: 2000-12-04									
; NUMBER OF SEQ ID NOS: 28									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 21									
; LENGTH: 2828									
; TYPE: PRT									
; ORGANISM: homo sapiens									
US-09-991-630-21									
Query Match 53.9%; Score 41; DB 3; Length 2828;									
Best Local Similarity 35.7%; Pred. No. 1.3e+03;									
Matches	5;	Conservative	7;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	IVLRPSRVHNSEEN 14							

; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
; LENGTH: 2828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1175

Query Match 53.9%; Score 41; DB 4; Length 2828;
Best Local Similarity 35.7%; Pred. No. 1.3e+03;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLPRSVHNSHEN 14
Db 2238 VMKPAKIEHKEEN 2251

RESULT 23

US-10-072-012-607
; Sequence 607, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gross, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 607
; LENGTH: 2828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-607

Query Match 53.9%; Score 41; DB 4; Length 2828;
Best Local Similarity 35.7%; Pred. No. 1.3e+03;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLPRSVHNSHEN 14
Db 2238 VMKPAKIEHKEEN 2251

RESULT 24

US-10-072-012-608
; Sequence 608, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gross, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767

RESULT 30
US-10-314-669-133
; Sequence 133, Application US/10314669

Query Match 52.6%; Score 40; DB 4; Length 90;
Best Local Similarity 64.3%; Pred. No. 44;
Matches 9: Conservative 2; Mismatches 1; Indels

QY 2 VLRP--SRVHSEE 13
|||||:|||||:
Db 64 VLRPILLRIHNSQ 77

RESULT 32

US-10-732-620-119
; Sequence 119, Application US/10732620
; Publication No. US20050032186A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Shin, Hyun-Chul
; APPLICANT: Kwon, Heung-Sun
; TITLE OF INVENTION: REGULATORY ZINC FINGER PROTEINS
; FILE REFERENCE: 12279-009001
; CURRENT APPLICATION NUMBER: US/10/732.620
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US 60/431,892
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-732-620-119

Query Match 52.6%; Score 40; DB 5; Length 90;
Best Local Similarity 64.3%; Pred. No. 44;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHSEE 13
|||||:|||||:
Db 64 VLRPILLRIHNSQ 77

RESULT 33

US-09-738-626-5581
; Sequence 5581, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKOYO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738.626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5581
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5581

Query Match 52.6%; Score 40; DB 3; Length 115;
Best Local Similarity 72.7%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 PSRVHNSERT 15
|||||:|||||:
Db 43 PLRVHNSERT 53

RESULT 34

US-10-314-669-221
; Sequence 221, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horiim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoon
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-221

Query Match 52.6%; Score 40; DB 4; Length 210;
Best Local Similarity 64.3%; Pred. No. 11e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHSEE 13
|||||:|||||:
Db 184 VLRPILLRIHNSQ 197

RESULT 35

US-10-669-861-221
; Sequence 221, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Lee, Yangsoon
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 10/314,669
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05

; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-221

Query Match 52.6%; Score 40; DB 4; Length 210;
Best Local Similarity 64.3%; Pred. No. 1.1e-02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHSEE 13
||||| :|||:
DB 184 VLRPILLRIHNSQ 197

RESULT 36
US-10-314-669-222
; Sequence 222, Application US/10314669
; Publication No. US20030194727A1

; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoo
; APPLICANT: Jang, Young-Soon

; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-222

Query Match 52.6%; Score 40; DB 4; Length 238;
Best Local Similarity 64.3%; Pred. No. 1.3e-02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHSEE 13
||||| :|||:
DB 212 VLRPILLRIHNSQ 225

RESULT 37
US-10-314-669-223
; Sequence 223, Application US/10314669
; Publication No. US20030194727A1

; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi

; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoo
; APPLICANT: Jang, Young-Soon
; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-223

Query Match 52.6%; Score 40; DB 4; Length 238;
Best Local Similarity 64.3%; Pred. No. 1.3e-02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVHSEE 13
||||| :|||:
DB 212 VLRPILLRIHNSQ 225

RESULT 38
US-10-314-669-224
; Sequence 224, Application US/10314669
; Publication No. US20030194727A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Seol, Wongi
; APPLICANT: Lee, Horim
; APPLICANT: Lee, Seong-il
; APPLICANT: Yang, Hyo-Young
; APPLICANT: Lee, Yangsoo
; APPLICANT: Jang, Young-Soon

; TITLE OF INVENTION: PHENOTYPIC SCREEN OF CHIMERIC PROTEINS
; FILE REFERENCE: 12279-007001
; CURRENT APPLICATION NUMBER: US/10/314,669
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-314-669-224

Query Match 52.6%; Score 40; DB 4; Length 238;

Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVNSE 13
||||| :||||:
Db 212 VLRPILLRIHNS 225

RESULT 39
US-10-669-861-222
; Sequence 222, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-222

Query Match 52.6%; Score 40; DB 4; Length 238;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVNSE 13
||||| :||||:
Db 212 VLRPILLRIHNS 225

RESULT 40
US-10-669-861-223
; Sequence 223, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 223
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-223

Query Match 52.6%; Score 40; DB 4; Length 238;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVNSE 13
||||| :||||:
Db 212 VLRPILLRIHNS 225

RESULT 41
US-10-669-861-224
; Sequence 224, Application US/10669861
; Publication No. US20040209277A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Dong-Ki
; APPLICANT: Kim, Jin-Soo
; TITLE OF INVENTION: DIFFERENTIATION PROTEINS
; FILE REFERENCE: 12279-007002
; CURRENT APPLICATION NUMBER: US/10/669,861
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 60/338,441
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/376,053
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/400,904
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/401,089
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid sequence
US-10-669-861-224

Query Match 52.6%; Score 40; DB 4; Length 238;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 2 VLRP--SRVNSE 13
||||| :||||:
Db 212 VLRPILLRIHNS 225

RESULT 42
US-10-425-114-72739
; Sequence 72739, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

Mon Mar 13 10:54:35 2006

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72739
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3587-232-A9_FLI.pep
US-10-425-114-72739

Query Match      52.6%; Score 40; DB 4; Length 313;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNS 11
Db 241 VMLRPLQIHNS 251

RESULT 43
US-10-437-963-157183
; Sequence 157183, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157183
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(671)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5677C.i.pep
US-10-437-963-157183

Query Match      52.6%; Score 40; DB 4; Length 671;
Best Local Similarity 64.3%; Pred. No. 4e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNS 15
Db 620 VLSPGAKNSENT 633

RESULT 44
US-10-149-310-106
; Sequence 106, Application US/10149310
; Publication No. US2004007039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir
; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
; TITLE OF INVENTION: Zinc Binuclear Cluster Proteins
; FILE REFERENCE: 14184-019US1
; CURRENT APPLICATION NUMBER: US/10/149,310
; CURRENT FILING DATE: 2003-02-19
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/29288
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/233,564
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-149-310-106

Query Match      52.6%; Score 40; DB 4; Length 836;
Best Local Similarity 64.3%; Pred. No. 5.1e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

Qy 2 VLRP--SRVHNS 13
Db 571 VLRPILLRIHNS 584

RESULT 45
US-11-097-143-40320
; Sequence 40320, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40320
; LENGTH: 975
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-40320

Query Match      52.6%; Score 40; DB 6; Length 975;
Best Local Similarity 54.5%; Pred. No. 6e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PSRVHNS 15
Db 84 PDRIHSDRNT 94

RESULT 46
US-10-898-141-17
; Sequence 17, Application US/10898141
; Publication No. US20050112699A1
; GENERAL INFORMATION:
; APPLICANT: SUTOVSKY, PETER
```

; APPLICANT: VAN LEYEN, KLAUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR EVALUATION AND MODULATION
; FILE REFERENCE: UYWO:0311US
; CURRENT APPLICATION NUMBER: US/10/898,141
; CURRENT FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: 60/489,871
; PRIOR FILING DATE: 2003-07-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 17
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Porcarius
US-10-898-141-17

Query Match 51.3%; Score 39; DB 5; Length 13;
Best Local Similarity 88.9%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LRPSRVHNS 11
Db 2 LRPSRVNS 10

RESULT 47
US-10-425-115-253664
; Sequence 253664, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253664
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(124)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MPT4577_162922C.1.pep
US-10-425-115-253664

Query Match 51.3%; Score 39; DB 4; Length 124;
Best Local Similarity 40.0%; Pred. No. 93;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSEENT 15
Db 49 IGIKPARIHENTXNT 63

RESULT 48
US-10-437-963-103361
; Sequence 103361, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103361
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100799C.1.pep
US-10-437-963-103361

Query Match 51.3%; Score 39; DB 4; Length 417;
Best Local Similarity 63.8%; Pred. No. 3.5e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
Db 33 LRPSRTHNTSE 43

RESULT 49
US-10-450-763-58555
; Sequence 58555, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 58555
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (119)..(128)
; OTHER INFORMATION: Zinc finger C3HC4 type (RING finger), proteins. domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00518, p-value=2.800e-1
; OTHER INFORMATION: raw score of 12.23
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (104)..(144)
; OTHER INFORMATION: Zinc finger, C3HC4 type (RING finger) domain identified by
; OTHER INFORMATION: Pfam, accession name zf-C3HC4, E-value=6.2e-13, Pfam score of 46.
US-10-450-763-58555

Query Match 51.3%; Score 39; DB 5; Length 421;
Best Local Similarity 72.7%; Pred. No. 3.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
Db 391 LRDSRIHVSE 401

RESULT 50
US-09-796-753-116
; Sequence 116, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:

APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,587
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 116
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-796-753-116

Query Match 51.3%; Score 39; DB 3; Length 497;
Best Local Similarity 58.3%; Pred. No. 4.3e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Oy 4 RPSRVHNSSEENT 15
Db 314 RPSRAHGREHNT 325

RESULT 51
US-10-437-963-172291
; Sequence 172291, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172291
; LENGTH: 946
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70441C.1.pep
US-10-437-963-172291

Query Match 51.3%; Score 39; DB 4; Length 946;
Best Local Similarity 63.6%; Pred. No. 8.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 2 VLPRSRVHNSSE 12
Db 74 ILRPADSHNSSE 84

RESULT 52
US-10-425-115-296047
; Sequence 296047, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 296047
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(60)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_33072C.1.pep
US-10-425-115-296047

Query Match 50.0%; Score 38; DB 4; Length 60;
Best Local Similarity 42.9%; Pred. No. 63;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Oy 1 IVLRPSRVHNSSEEN 14
Db 33 MILRIAIAHRSEDN 46

RESULT 53


```

; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35808
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011307.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.6
; OTHER INFORMATION: EST HUMAN HIT: AW937782.1, EVALUATE 7.00e-79
; OTHER INFORMATION: SWISSPROT HIT: O35136, EVALUATE 8.00e-12
; US-09-864-761-35808

Query Match 50.0%; Score 38; DB 3; Length 183;
Best Local Similarity 28.6%; Pred. No. 2.1e+02;
Matches 4; Conservative 8; Mismatches 2; Indels

QY 1 IVLRPSRVHNSSEN 14
   :::|:::|:::
DB 45 VMKPKAIEHKEEN 58

RESULT 56
US-09-815-242-5560
; Sequence 5560, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haebebeck, Robert
; APPLICANT: Ohlsen, Karl L.

```

APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5560
LENGTH: 232
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5560

Query Match 50.0%; Score 38; DB 3; Length 232;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHSEE 13
Db 129 RPSQDNSEE 138

RESULT 57
US-09-815-242-12569
Sequence 12569, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohleen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12569
LENGTH: 232
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12569

Query Match 50.0%; Score 38; DB 3; Length 232;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RPSRVHSEE 13
Db 129 RPSQDNSEE 138

RESULT 58
US-10-767-701-32297
Sequence 32297, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 32297
LENGTH: 235
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C123887_1.pep
US-10-767-701-32297

Query Match 50.0%; Score 38; DB 4; Length 235;
Best Local Similarity 46.2%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHSEE 13
Db 125 VVVEPSRTHRDXE 137

RESULT 59
US-10-437-963-204794
Sequence 204794, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 204794
LENGTH: 242
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_99846C.1.pep
US-10-437-963-204794

Query Match 50.0%; Score 38; DB 4; Length 242;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PSRVHNSSEN 14
Db 11 PSRLHSTESN 20

RESULT 60

US-10-767-701-35712
; Sequence 35712, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 35712
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C72553_1.pep
US-10-767-701-35712

Query Match 50.0%; Score 38; DB 4; Length 303;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRFPSRVHNSSEN 15
Db 99 VIRKFRGHNSSEVNS 112

Search completed: March 11, 2006, 12:32:19
Job time : 126.632 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:27:37 ; Search time 13.8158 Seconds
(without alignments)
30.221 Million cell updates/sec

Title: US-10-774-242A-1

Perfect score: 76

Sequence: 1 IVLPRSRVHSEENT 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications AA New:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	760	7	US-11-186-284-55
2	41	53.9	2828	7	US-11-080-991-54
3	41	53.9	2828	7	US-11-186-284-49
4	40	52.6	331	7	US-11-096-568A-21152
5	40	52.6	377	7	US-11-096-568A-21150
6	40	52.6	532	7	US-11-087-099-7794
7	39	51.3	917	7	US-11-096-568A-32095
8	39	51.3	949	7	US-11-096-568A-32094
9	38	50.0	252	7	US-11-087-099-11054
10	37	48.7	507	7	US-11-096-568A-28365
11	37	48.7	632	7	US-11-096-568A-28364
12	37	48.7	732	7	US-11-096-568A-28363
13	37	48.7	2479	7	US-11-087-099-2391
14	36	47.4	420	6	US-10-935-494-3
15	36	47.4	821	7	US-11-124-367A-473
16	36	47.4	1394	6	US-10-935-494-31
17	35	46.1	205	7	US-11-087-099-835
18	35	46.1	347	7	US-11-234-786-590
19	35	46.1	542	7	US-11-096-568A-11819
20	35	46.1	555	7	US-11-096-568A-11818
21	35	46.1	838	7	US-11-087-099-829
22	34	44.7	181	7	US-11-087-099-1383
23	34	44.7	363	7	US-11-096-568A-5328
24	34	44.7	377	7	US-11-124-368A-209
25	34	44.7	377	7	US-11-124-368A-210

486	7	US-11-087-099-4550	Sequence 4550, Ap
581	7	US-11-057-231-153	Sequence 153, App
581	7	US-11-054-281-22	Sequence 22, Appl
581	7	US-11-054-281-91	Sequence 91, Appl
588	6	US-10-793-626-3328	Sequence 3328, Ap
593	7	US-11-087-099-3671	Sequence 3671, Ap
594	7	US-11-087-099-9040	Sequence 9040, Ap
594	7	US-10-055-877-157	Sequence 157, App
757	6	US-10-821-234-1592	Sequence 1592, Ap
782	6	US-11-177-506-38	Sequence 38, Appl
782	7	US-11-177-506-38	Sequence 5, Appl
1311	6	US-10-509-422-5	Sequence 3, Appl
1311	6	US-11-179-624-3	Sequence 115, App
1980	6	US-10-884-730-115	Sequence 115, App
1980	6	US-11-087-099-287	Sequence 287, App
173	7	US-11-096-568A-27299	Sequence 27299, A
180	7	US-11-096-568A-328	Sequence 328, App
184	7	US-11-087-099-5689	Sequence 5689, Ap
231	7	US-11-096-568A-327	Sequence 327, App
231	7	US-11-096-568A-329	Sequence 329, App
231	7	US-10-793-626-1894	Sequence 1894, Ap
241	6	US-10-884-730-114	Sequence 114, App
266	6	US-11-087-099-287	Sequence 287, App
287	7	US-11-096-568A-19351	Sequence 19351, A
289	7	US-11-096-568A-19350	Sequence 19350, A
295	7	US-11-024-959-267	Sequence 267, App
313	7	US-11-096-568A-19349	Sequence 19349, A
315	7	US-11-096-568A-12602	Sequence 12602, A
339	6	US-10-467-657-5522	Sequence 5522, Ap
387	6	US-11-069-642-92	Sequence 92, Appl
437	7	US-11-069-642-47	Sequence 47, Appl
437	7	US-11-069-642-49	Sequence 49, Appl
438	7	US-11-069-642-51	Sequence 51, Appl
438	7	US-11-069-642-53	Sequence 53, Appl
438	7	US-11-069-642-55	Sequence 55, Appl
438	7	US-11-069-642-57	Sequence 57, Appl
438	7	US-11-069-642-59	Sequence 59, Appl
438	7	US-11-069-642-61	Sequence 61, Appl
438	7	US-11-069-642-63	Sequence 63, Appl
438	7	US-11-096-568A-31804	Sequence 31804, A
546	7	US-11-096-568A-31803	Sequence 31803, A
567	7	US-10-793-626-2844	Sequence 2844, Ap
596	6	US-11-057-047-2	Sequence 2, Appl
628	6	US-11-057-047-1	Sequence 1, Appl
739	7	US-11-057-047-2	Sequence 2, Appl
764	7	US-11-057-047-1	Sequence 1, Appl
779	6	US-11-096-568A-32062	Sequence 32062, A
798	6	US-10-821-234-1034	Sequence 1034, Ap
916	6	US-11-096-568A-32061	Sequence 32061, A
939	7	US-11-096-568A-32060	Sequence 32060, A
1053	6	US-10-330-773-34	Sequence 34, Appl
1063	6	US-10-330-773-31	Sequence 31, Appl
1085	6	US-10-330-773-36	Sequence 36, Appl
1087	7	US-11-102-978-2	Sequence 2, Appl
1307	6	US-10-995-561-711	Sequence 711, App
1344	6	US-11-072-512-2452	Sequence 2452, Ap
1495	6	US-10-453-372-1002	Sequence 1002, Ap
5636	7	US-11-065-695-20	Sequence 20, Appl
191	7	US-11-096-568A-3655	Sequence 3655, Ap
325	7	US-11-087-099-1363	Sequence 1363, Ap
325	7	US-11-153-667-35	Sequence 35, Appl
89	7	US-11-072-512-3817	Sequence 3817, Ap
105	7	US-11-096-568A-568	Sequence 568, App
118	7	US-11-096-568A-5272	Sequence 5272, Ap
150	7	US-11-096-568A-567	Sequence 567, App
156	7	US-11-096-568A-5271	Sequence 5271, Ap
161	7	US-11-096-568A-5270	Sequence 5270, Ap
173	7	US-11-096-568A-20264	Sequence 20264, A
191	7	US-11-096-568A-32542	Sequence 32542, A
239	7	US-10-330-773-758	Sequence 758, App
273	6	US-11-096-568A-27927	Sequence 27927, A
277	7	US-10-467-657-1326	Sequence 1326, Ap
280	6	US-11-102-497-6	Sequence 6, Appl
280	7	US-11-096-568A-5326	Sequence 5326, Ap
282	7	US-11-096-568A-27926	Sequence 27926, A
296	7	US-11-054-281-139	Sequence 139, App
304	7		

99	32	42.1	309	7	US-11-096-568A-5325	Sequence 5325, Ap	172	31	40.8	245	6	US-10-467-657-1570	Sequence 1570, Ap
100	32	42.1	311	7	US-11-096-568A-13492	Sequence 13492, A	173	31	40.8	263	7	US-11-087-099-3093	Sequence 3093, Ap
101	32	42.1	313	7	US-11-087-099-7652	Sequence 7652, Ap	174	31	40.8	265	7	US-11-096-568A-21917	Sequence 21917, A
102	32	42.1	316	7	US-11-087-099-7594	Sequence 7594, Ap	175	31	40.8	269	7	US-11-096-568A-26985	Sequence 26985, A
103	32	42.1	322	7	US-11-087-099-5811	Sequence 5811, Ap	176	31	40.8	273	7	US-11-096-568A-13827	Sequence 13827, A
104	32	42.1	322	7	US-11-087-099-12315	Sequence 12315, A	177	31	40.8	275	7	US-11-096-568A-24634	Sequence 24634, A
105	32	42.1	326	7	US-11-096-568A-27925	Sequence 27925, A	178	31	40.8	293	7	US-11-096-568A-16425	Sequence 16425, A
106	32	42.1	327	7	US-11-098-586-11321	Sequence 11321, A	179	31	40.8	294	7	US-11-096-568A-9328	Sequence 9328, A
107	32	42.1	335	7	US-11-096-568A-5324	Sequence 5324, Ap	180	31	40.8	294	7	US-11-096-568A-9330	Sequence 9330, Ap
108	32	42.1	336	7	US-11-096-568A-32541	Sequence 32541, A	181	31	40.8	299	6	US-10-511-538-26	Sequence 26, Appl
109	32	42.1	338	7	US-11-096-568A-32540	Sequence 32540, A	182	31	40.8	299	7	US-11-024-959-359	Sequence 359, Appl
110	32	42.1	343	7	US-11-087-099-9361	Sequence 9361, A	183	31	40.8	315	7	US-11-096-568A-24633	Sequence 24633, A
111	32	42.1	343	7	US-11-087-099-11908	Sequence 11908, A	184	31	40.8	320	7	US-11-198-728-15	Sequence 15, Appl
112	32	42.1	363	7	US-11-087-099-5838	Sequence 5838, Ap	185	31	40.8	322	7	US-11-050-857-96	Sequence 96, Appl
113	32	42.1	364	7	US-11-087-099-3649	Sequence 3649, Ap	186	31	40.8	327	7	US-11-098-686-11417	Sequence 11417, A
114	32	42.1	376	7	US-11-182-752-2	Sequence 2, Appl	187	31	40.8	327	7	US-11-096-568A-9327	Sequence 9327, Ap
115	32	42.1	386	6	US-10-131-826A-340	Sequence 340, App	188	31	40.8	332	7	US-11-098-686-11282	Sequence 11282, A
116	32	42.1	386	6	US-10-973-115B-340	Sequence 340, App	189	31	40.8	337	7	US-11-087-099-10558	Sequence 10558, A
117	32	42.1	386	7	US-11-185-878-2	Sequence 2, Appl	190	31	40.8	338	6	US-10-921-793-54	Sequence 54, Appl
118	32	42.1	386	7	US-11-099-135-1	Sequence 1, Appl	191	31	40.8	338	6	US-10-931-198-54	Sequence 54, Appl
119	32	42.1	407	7	US-11-087-099-9116	Sequence 9116, Ap	192	31	40.8	339	6	US-10-467-657-3018	Sequence 3018, Ap
120	32	42.1	417	7	US-11-182-752-4	Sequence 4, Appl	193	31	40.8	343	7	US-11-096-568A-7482	Sequence 7482, Ap
121	32	42.1	417	7	US-11-054-281-40	Sequence 40, Appl	194	31	40.8	344	7	US-11-096-568A-7481	Sequence 7481, Ap
122	32	42.1	417	7	US-11-054-281-136	Sequence 136, Appl	195	31	40.8	353	7	US-11-096-568A-24793	Sequence 24793, A
123	32	42.1	459	7	US-11-087-099-3410	Sequence 3410, Ap	196	31	40.8	357	7	US-11-087-099-6967	Sequence 6967, Ap
124	32	42.1	459	7	US-11-087-099-4535	Sequence 4535, Ap	197	31	40.8	358	7	US-11-096-568A-24632	Sequence 24632, A
125	32	42.1	462	7	US-11-087-099-8690	Sequence 8690, Ap	198	31	40.8	363	7	US-11-087-099-9335	Sequence 9335, Ap
126	32	42.1	469	7	US-11-024-959-271	Sequence 271, App	199	31	40.8	364	7	US-11-087-099-11517	Sequence 11517, A
127	32	42.1	476	7	US-11-096-568A-13491	Sequence 13491, A	200	31	40.8	364	7	US-11-096-568A-7480	Sequence 7480, Ap
128	32	42.1	491	7	US-11-096-568A-13490	Sequence 13490, A	201	31	40.8	365	7	US-11-087-099-2798	Sequence 2798, Ap
129	32	42.1	516	6	US-10-467-657-7038	Sequence 7038, Ap	202	31	40.8	373	7	US-11-050-857-98	Sequence 98, Appl
130	32	42.1	541	7	US-11-098-686-11058	Sequence 11058, A	203	31	40.8	380	7	US-11-144-236-1	Sequence 1, Appl
131	32	42.1	587	7	US-11-087-099-11657	Sequence 11657, A	204	31	40.8	382	7	US-11-050-857-97	Sequence 97, Appl
132	32	42.1	690	6	US-10-131-826A-306	Sequence 306, App	205	31	40.8	391	6	US-10-467-657-2524	Sequence 2524, A
133	32	42.1	690	6	US-10-973-115B-306	Sequence 306, App	206	31	40.8	392	7	US-11-096-568A-20292	Sequence 2, A
134	32	42.1	740	7	US-11-096-568A-24714	Sequence 24714, A	207	31	40.8	401	6	US-10-510-876-2	Sequence 2, Appl
135	32	42.1	777	7	US-11-096-568A-24713	Sequence 24713, A	208	31	40.8	401	6	US-10-510-876-4	Sequence 4, Appl
136	32	42.1	790	7	US-11-096-568A-24712	Sequence 24712, A	209	31	40.8	401	6	US-10-921-793-52	Sequence 52, Appl
137	32	42.1	840	7	US-11-108-172-1102	Sequence 1102, Ap	210	31	40.8	401	6	US-10-931-198-52	Sequence 52, Appl
138	32	42.1	1037	6	US-10-392-234A-36	Sequence 36, Appl	211	31	40.8	404	7	US-11-072-512-3621	Sequence 3621, Ap
139	32	42.1	1928	6	US-10-480-330-30	Sequence 30, Appl	212	31	40.8	419	7	US-11-087-099-4639	Sequence 4639, Ap
140	32	42.1	1965	6	US-10-480-330-2	Sequence 4, Appl	213	31	40.8	421	7	US-11-055-822-658	Sequence 658, App
141	32	42.1	1966	6	US-10-480-330-6	Sequence 6, Appl	214	31	40.8	424	7	US-11-229-371-98	Sequence 98, Appl
142	32	42.1	1966	6	US-10-480-330-8	Sequence 8, Appl	215	31	40.8	424	7	US-11-229-371-110	Sequence 110, App
143	32	42.1	1966	6	US-10-480-330-10	Sequence 10, Appl	216	31	40.8	424	7	US-11-229-371-117	Sequence 117, App
144	32	42.1	1966	6	US-10-480-330-12	Sequence 12, Appl	217	31	40.8	424	7	US-11-229-371-119	Sequence 119, App
145	32	42.1	1966	6	US-10-480-330-14	Sequence 14, Appl	218	31	40.8	424	7	US-11-228-923-98	Sequence 98, Appl
146	32	42.1	1966	6	US-10-480-330-16	Sequence 16, Appl	219	31	40.8	424	7	US-11-228-923-110	Sequence 110, App
147	32	42.1	1966	6	US-10-480-330-18	Sequence 18, Appl	220	31	40.8	424	7	US-11-228-923-117	Sequence 117, App
148	32	42.1	1966	6	US-10-480-330-20	Sequence 20, Appl	221	31	40.8	424	7	US-11-228-923-119	Sequence 119, App
149	32	42.1	1966	6	US-10-480-330-22	Sequence 22, Appl	222	31	40.8	424	7	US-11-228-923-125	Sequence 125, App
150	32	42.1	1966	6	US-10-480-330-24	Sequence 24, Appl	223	31	40.8	424	7	US-11-228-923-125	Sequence 125, App
151	32	42.1	1966	6	US-10-480-330-26	Sequence 26, Appl	224	31	40.8	424	7	US-11-228-875-98	Sequence 98, Appl
152	32	42.1	1966	6	US-10-480-330-28	Sequence 28, Appl	225	31	40.8	424	7	US-11-228-875-110	Sequence 110, App
153	32	42.1	1966	6	US-10-480-330-28	Sequence 28, Appl	226	31	40.8	424	7	US-11-228-875-117	Sequence 117, App
154	32	42.1	1970	6	US-10-821-234-1641	Sequence 1641, Ap	227	31	40.8	424	7	US-11-228-875-119	Sequence 119, App
155	32	42.1	2098	6	US-10-055-877-253	Sequence 253, App	228	31	40.8	424	7	US-11-228-875-125	Sequence 125, App
156	32	42.1	2326	7	US-11-126-313-37	Sequence 37, Appl	229	31	40.8	430	7	US-11-096-568A-12991	Sequence 12991, A
157	31	40.8	81	7	US-11-096-568A-9799	Sequence 9799, Ap	230	31	40.8	430	7	US-11-096-568A-12991	Sequence 32239, A
158	31	40.8	101	6	US-10-467-657-6572	Sequence 6572, Ap	231	31	40.8	448	7	US-11-037-243-103	Sequence 103, App
159	31	40.8	101	6	US-10-467-657-7642	Sequence 7642, Ap	232	31	40.8	448	7	US-11-096-568A-12990	Sequence 12990, A
160	31	40.8	101	7	US-11-103-957-53	Sequence 53, Appl	233	31	40.8	457	7	US-11-087-099-5823	Sequence 5823, Ap
161	31	40.8	109	6	US-10-821-234-1175	Sequence 1175, Ap	234	31	40.8	458	6	US-10-763-712A-46	Sequence 46, Appl
162	31	40.8	146	6	US-10-467-657-2680	Sequence 2680, Ap	235	31	40.8	464	6	US-10-763-712A-3	Sequence 3, Appl
163	31	40.8	151	7	US-11-072-512-3705	Sequence 3705, Ap	236	31	40.8	470	7	US-11-096-568A-32238	Sequence 32238, A
164	31	40.8	154	7	US-11-072-512-1992	Sequence 1992, Ap	237	31	40.8	479	7	US-11-096-568A-12989	Sequence 12989, A
165	31	40.8	167	7	US-11-096-568A-9329	Sequence 9329, Ap	238	31	40.8	479	7	US-11-087-099-239	Sequence 239, A
166	31	40.8	209	6	US-10-995-561-736	Sequence 736, App	239	31	40.8	512	6	US-10-517-939-236	Sequence 236, App
167	31	40.8	210	6	US-10-467-657-4386	Sequence 4386, Ap	240	31	40.8	520	6	US-10-467-657-3948	Sequence 3948, Ap
168	31	40.8	221	6	US-10-510-386-210	Sequence 210, App	241	31	40.8	537	7	US-11-144-236-6	Sequence 6, Appl
169	31	40.8	228	7	US-11-087-099-376	Sequence 376, App	242	31	40.8	548	7	US-11-096-568A-32237	Sequence 32237, A
170	31	40.8	233	7	US-11-050-857-99	Sequence 99, Appl	243	31	40.8	555	7	US-11-024-959-270	Sequence 270, App
171	31	40.8	241	7	US-11-096-568A-13828	Sequence 13828, A	244	31	40.8				

245	31	40.8	577	7	US-11-072-512-2569	Sequence 2569, Ap	318	30	39.5	89	6	US-10-884-730-27	Sequence 27, Appl
246	31	40.8	589	6	US-11-096-568A-33658	Sequence 33658, A	319	30	39.5	89	6	US-10-884-730-30	Sequence 30, Appl
247	31	40.8	619	6	US-10-763-712A-35	Sequence 35, Appl	320	30	39.5	89	6	US-10-884-730-31	Sequence 31, Appl
248	31	40.8	626	7	US-11-010-748A-1	Sequence 1, Appl	321	30	39.5	89	6	US-10-884-730-137	Sequence 137, App
249	31	40.8	643	7	US-11-096-568A-34196	Sequence 34196, A	322	30	39.5	89	6	US-10-884-730-176	Sequence 176, App
250	31	40.8	656	7	US-11-096-568A-34195	Sequence 34195, A	323	30	39.5	89	6	US-10-884-730-182	Sequence 182, App
251	31	40.8	660	7	US-11-186-284-125	Sequence 125, App	324	30	39.5	89	6	US-10-884-730-188	Sequence 188, App
252	31	40.8	665	7	US-11-096-568A-34194	Sequence 34194, A	325	30	39.5	89	6	US-10-884-730-197	Sequence 197, App
253	31	40.8	670	6	US-10-821-234-879	Sequence 879, App	326	30	39.5	89	6	US-10-884-730-205	Sequence 205, App
254	31	40.8	708	6	US-10-821-234-917	Sequence 917, App	327	30	39.5	89	6	US-10-884-730-206	Sequence 206, App
255	31	40.8	709	7	US-11-096-568A-29889	Sequence 29889, A	328	30	39.5	89	6	US-10-884-730-208	Sequence 208, App
256	31	40.8	778	7	US-11-096-568A-29888	Sequence 29888, A	329	30	39.5	89	6	US-10-884-730-210	Sequence 210, App
257	31	40.8	803	7	US-11-096-568A-29887	Sequence 29887, A	330	30	39.5	89	6	US-10-884-730-211	Sequence 211, App
258	31	40.8	814	7	US-11-096-568A-28352	Sequence 28352, A	331	30	39.5	89	6	US-10-884-730-213	Sequence 213, App
259	31	40.8	816	7	US-11-096-568A-28351	Sequence 28351, A	332	30	39.5	89	6	US-10-884-730-218	Sequence 218, App
260	31	40.8	821	6	US-10-912-580-9	Sequence 9, Appl	333	30	39.5	89	6	US-10-884-730-220	Sequence 220, App
261	31	40.8	821	6	US-10-912-582-3	Sequence 3, Appl	334	30	39.5	89	6	US-10-884-730-221	Sequence 221, App
262	31	40.8	821	6	US-11-050-857-964	Sequence 964, App	335	30	39.5	89	6	US-10-884-730-222	Sequence 222, App
263	31	40.8	847	7	US-11-087-099-3302	Sequence 3302, Ap	336	30	39.5	89	6	US-10-884-730-256	Sequence 256, App
264	31	40.8	855	7	US-11-096-568A-29960	Sequence 29960, A	337	30	39.5	89	6	US-10-884-730-258	Sequence 258, App
265	31	40.8	876	7	US-11-087-099-10193	Sequence 10193, A	338	30	39.5	89	6	US-10-884-730-248	Sequence 248, App
266	31	40.8	882	6	US-10-912-580-8	Sequence 8, Appl	339	30	39.5	90	6	US-11-096-568A-9319	Sequence 9319, Ap
267	31	40.8	882	6	US-10-912-582-2	Sequence 2, Appl	340	30	39.5	94	6	US-10-884-730-231	Sequence 231, App
268	31	40.8	882	7	US-11-050-857-95	Sequence 95, Appl	341	30	39.5	94	6	US-10-884-730-232	Sequence 232, App
269	31	40.8	897	7	US-11-096-568A-29959	Sequence 29959, A	342	30	39.5	95	6	US-11-159-667-30	Sequence 30, Appl
270	31	40.8	901	7	US-11-050-857-963	Sequence 963, App	343	30	39.5	95	6	US-10-884-730-253	Sequence 253, App
271	31	40.8	911	7	US-11-096-568A-29958	Sequence 29958, A	344	30	39.5	100	6	US-10-884-730-13	Sequence 13, Appl
272	31	40.8	926	7	US-11-087-099-345	Sequence 345, App	345	30	39.5	102	6	US-10-884-730-19	Sequence 19, Appl
273	31	40.8	986	7	US-11-203-251A-80	Sequence 80, Appl	346	30	39.5	102	6	US-10-884-730-20	Sequence 20, Appl
274	31	40.8	1009	6	US-10-514-531-11	Sequence 11, Appl	347	30	39.5	102	6	US-11-086-568A-12851	Sequence 12851, A
275	31	40.8	1011	6	US-10-330-773-924	Sequence 924, App	348	30	39.5	103	6	US-10-467-657-8753	Sequence 8753, Ap
276	31	40.8	1148	7	US-11-110-082-29	Sequence 29, Appl	349	30	39.5	103	6	US-11-084-554-226	Sequence 226, App
277	31	40.8	1166	6	US-10-501-035-205	Sequence 205, App	350	30	39.5	104	7	US-11-136-250-226	Sequence 226, App
278	31	40.8	1207	6	US-10-821-234-1109	Sequence 1109, App	351	30	39.5	104	7	US-11-072-512-3422	Sequence 3422, Ap
279	31	40.8	3635	7	US-11-013-711-47	Sequence 47, Appl	352	30	39.5	110	7	US-11-096-568A-18617	Sequence 18617, A
280	30	39.5	74	6	US-10-884-730-241	Sequence 241, App	353	30	39.5	110	7	US-11-087-099-9510	Sequence 9510, Ap
281	30	39.5	76	6	US-10-884-730-183	Sequence 183, App	354	30	39.5	125	7	US-11-096-568A-13902	Sequence 13902, A
282	30	39.5	77	6	US-10-884-730-129	Sequence 129, App	355	30	39.5	127	7	US-11-096-568A-12849	Sequence 12849, A
283	30	39.5	77	6	US-11-159-667-3	Sequence 3, Appl	356	30	39.5	130	7	US-11-096-568A-13901	Sequence 13901, A
284	30	39.5	77	7	US-11-159-667-28	Sequence 28, Appl	357	30	39.5	136	7	US-11-096-568A-14105	Sequence 14105, A
285	30	39.5	77	7	US-11-159-667-40	Sequence 40, Appl	358	30	39.5	137	7	US-11-123-896-374	Sequence 374, App
286	30	39.5	78	6	US-10-884-730-242	Sequence 242, App	359	30	39.5	154	7	US-11-096-568A-14104	Sequence 14104, A
287	30	39.5	80	6	US-10-884-730-24	Sequence 24, Appl	360	30	39.5	158	7	US-11-098-686-10384	Sequence 10384, A
288	30	39.5	80	6	US-10-884-730-250	Sequence 250, App	361	30	39.5	160	7	US-11-087-099-10347	Sequence 10347, A
289	30	39.5	81	6	US-10-884-730-28	Sequence 28, Appl	362	30	39.5	161	7	US-11-087-099-5829	Sequence 5829, Ap
290	30	39.5	81	6	US-10-884-730-162	Sequence 162, App	363	30	39.5	168	7	US-11-072-512-2540	Sequence 2540, Ap
291	30	39.5	81	6	US-10-884-730-313	Sequence 313, App	364	30	39.5	179	7	US-11-086-568A-9456	Sequence 9456, Ap
292	30	39.5	82	6	US-10-884-730-195	Sequence 195, App	365	30	39.5	193	7	US-11-096-568A-18616	Sequence 18616, A
293	30	39.5	82	6	US-10-884-730-207	Sequence 207, App	366	30	39.5	194	7	US-11-064-774A-125	Sequence 5, Appl
294	30	39.5	84	6	US-10-884-730-134	Sequence 134, App	367	30	39.5	196	7	US-11-211-724-3	Sequence 3, Appl
295	30	39.5	84	6	US-10-884-730-259	Sequence 259, App	368	30	39.5	196	7	US-11-129-076-2	Sequence 2, Appl
296	30	39.5	84	6	US-11-140-284-12	Sequence 12, Appl	369	30	39.5	196	7	US-11-233-119-5	Sequence 5, Appl
297	30	39.5	85	6	US-10-884-730-14	Sequence 14, Appl	370	30	39.5	196	7	US-11-075-047A-99	Sequence 99, Appl
298	30	39.5	85	6	US-10-884-730-141	Sequence 141, App	371	30	39.5	196	7	US-11-096-568A-15843	Sequence 15843, A
299	30	39.5	85	6	US-10-884-730-187	Sequence 187, App	372	30	39.5	200	7	US-11-129-076-1	Sequence 1, Appl
300	30	39.5	86	6	US-10-884-730-18	Sequence 18, Appl	373	30	39.5	204	7	US-10-467-657-2020	Sequence 2020, Ap
301	30	39.5	86	6	US-10-884-730-247	Sequence 247, App	374	30	39.5	210	6	US-11-075-512-2179	Sequence 2179, Ap
302	30	39.5	86	6	US-10-884-730-282	Sequence 282, App	375	30	39.5	211	6	US-11-096-568A-24674	Sequence 24674, A
303	30	39.5	87	6	US-10-884-730-11	Sequence 11, Appl	376	30	39.5	225	7	US-11-096-568A-15842	Sequence 15842, A
304	30	39.5	87	6	US-10-884-730-219	Sequence 219, App	377	30	39.5	227	6	US-10-884-730-10	Sequence 10, Appl
305	30	39.5	87	6	US-10-884-730-238	Sequence 238, App	378	30	39.5	237	6	US-10-884-730-12	Sequence 12, Appl
306	30	39.5	88	6	US-10-884-730-32	Sequence 32, Appl	379	30	39.5	237	6	US-10-884-730-175	Sequence 175, App
307	30	39.5	88	6	US-10-884-730-178	Sequence 178, App	380	30	39.5	237	6	US-10-884-730-177	Sequence 177, App
308	30	39.5	88	6	US-10-884-730-199	Sequence 199, App	381	30	39.5	240	6	US-10-467-657-998	Sequence 998, App
309	30	39.5	88	6	US-10-884-730-214	Sequence 214, App	382	30	39.5	240	6	US-11-096-568A-27209	Sequence 27209, A
310	30	39.5	88	6	US-10-884-730-235	Sequence 235, App	383	30	39.5	243	7	US-11-087-099-663	Sequence 663, App
311	30	39.5	88	6	US-10-884-730-246	Sequence 246, App	384	30	39.5	245	7	US-11-096-568A-19534	Sequence 19534, A
312	30	39.5	88	6	US-10-884-730-305	Sequence 305, App	385	30	39.5				
313	30	39.5	89	6	US-10-884-730-16	Sequence 16, Appl	386	30	39.5				
314	30	39.5	89	6	US-10-884-730-17	Sequence 17, Appl	387	30	39.5				
315	30	39.5	89	6	US-10-884-730-23	Sequence 23, Appl	388	30	39.5				
316	30	39.5	89	6	US-10-884-730-25	Sequence 25, Appl	389	30	39.5				
317	30	39.5	89	6	US-10-884-730-26	Sequence 26, Appl	390	30	39.5				

391	30	39.5	247	7	US-11-096-568A-27676	Sequence 27676, A	464	30	39.5	757	7	US-11-096-568A-28588	Sequence 28588, A
392	30	39.5	248	7	US-11-096-568A-9411	Sequence 9411, A	465	30	39.5	759	6	US-10-467-657-2722	Sequence 2722, A
393	30	39.5	249	7	US-11-096-568A-27208	Sequence 27208, A	466	30	39.5	794	6	US-10-485-517-355	Sequence 355, A
394	30	39.5	251	7	US-11-096-568A-24673	Sequence 24673, A	467	30	39.5	822	7	US-11-183-567A-2	Sequence 2, Appli
395	30	39.5	255	7	US-11-096-568A-9861	Sequence 9861, A	468	30	39.5	851	7	US-11-096-568A-29030	Sequence 29030, A
396	30	39.5	257	7	US-11-096-568A-11406	Sequence 11406, A	469	30	39.5	852	7	US-11-077-550-145	Sequence 145, A
397	30	39.5	257	7	US-11-096-568A-9860	Sequence 9860, A	470	30	39.5	879	7	US-11-077-550-143	Sequence 143, A
398	30	39.5	263	7	US-11-096-568A-27207	Sequence 27207, A	471	30	39.5	887	7	US-11-077-550-147	Sequence 147, A
399	30	39.5	266	7	US-11-096-568A-21549	Sequence 21549, A	472	30	39.5	900	7	US-11-182-016-37	Sequence 37, Appl
400	30	39.5	267	6	US-10-467-657-2038	Sequence 2038, A	473	30	39.5	938	7	US-11-072-512-3855	Sequence 3855, A
401	30	39.5	267	6	US-10-467-657-2038	Sequence 2038, A	474	30	39.5	955	7	US-11-096-568A-23141	Sequence 23141, A
402	30	39.5	281	7	US-11-087-099-1752	Sequence 1752, A	475	30	39.5	989	6	US-10-793-626-2594	Sequence 2594, A
403	30	39.5	281	7	US-11-096-568A-27675	Sequence 27675, A	476	30	39.5	1015	6	US-11-096-568A-29140	Sequence 29140, A
404	30	39.5	282	7	US-11-096-568A-9410	Sequence 9410, A	477	30	39.5	1018	7	US-11-067-121-17	Sequence 17, Appl
405	30	39.5	288	7	US-11-087-099-7007	Sequence 7007, A	478	30	39.5	1036	6	US-10-131-826A-142	Sequence 142, A
406	30	39.5	294	6	US-10-858-730-118	Sequence 118, A	479	30	39.5	1036	6	US-10-973-1158-142	Sequence 142, A
407	30	39.5	306	7	US-11-096-568A-27674	Sequence 27674, A	480	30	39.5	1036	6	US-11-096-568A-28139	Sequence 28139, A
408	30	39.5	310	7	US-11-087-099-5412	Sequence 5412, A	481	30	39.5	1096	6	US-10-995-561-710	Sequence 710, A
409	30	39.5	311	7	US-11-096-568A-24672	Sequence 24672, A	482	30	39.5	1272	6	US-10-501-035-313	Sequence 313, A
410	30	39.5	324	7	US-11-096-568A-9409	Sequence 9409, A	483	30	39.5	1299	7	US-11-169-041-231	Sequence 231, A
411	30	39.5	338	6	US-10-467-657-4976	Sequence 4976, A	484	30	39.5	1304	6	US-10-821-234-1648	Sequence 1648, A
412	30	39.5	340	7	US-11-055-822-270	Sequence 270, A	485	30	39.5	1315	7	US-11-077-550-141	Sequence 141, A
413	30	39.5	340	7	US-11-087-099-8225	Sequence 8225, A	486	30	39.5	1330	6	US-10-453-372-260	Sequence 260, A
414	30	39.5	341	6	US-10-821-234-1628	Sequence 1628, A	487	30	39.5	1440	7	US-11-096-568A-28130	Sequence 28130, A
415	30	39.5	341	7	US-11-098-686-11432	Sequence 11432, A	488	30	39.5	1441	7	US-11-096-568A-28129	Sequence 28129, A
416	30	39.5	352	7	US-11-087-099-3715	Sequence 3715, A	489	30	39.5	1450	6	US-10-485-517-152	Sequence 152, A
417	30	39.5	372	7	US-11-096-568A-18959	Sequence 18959, A	490	30	39.5	1490	7	US-11-096-568A-28128	Sequence 28128, A
418	30	39.5	377	7	US-11-096-568A-18958	Sequence 18958, A	491	30	39.5	1501	6	US-10-793-626-2850	Sequence 2850, A
419	30	39.5	381	7	US-11-096-568A-14664	Sequence 14664, A	492	30	39.5	1542	6	US-10-453-372-258	Sequence 258, A
420	30	39.5	382	5	US-09-978-360A-425	Sequence 425, A	493	30	39.5	1542	6	US-10-453-372-266	Sequence 266, A
421	30	39.5	383	7	US-11-096-568A-19533	Sequence 19533, A	494	30	39.5	1542	6	US-10-453-372-280	Sequence 280, A
422	30	39.5	391	7	US-11-087-099-5368	Sequence 5368, A	495	30	39.5	1551	7	US-11-087-099-10366	Sequence 10366, A
423	30	39.5	391	7	US-11-096-568A-19532	Sequence 19532, A	496	30	39.5	1575	6	US-10-501-035-228	Sequence 228, A
424	30	39.5	393	7	US-11-194-246-309	Sequence 309, A	497	30	39.5	1614	6	US-10-821-234-903	Sequence 903, A
425	30	39.5	397	7	US-11-096-568A-30094	Sequence 30094, A	498	30	39.5	1669	6	US-10-330-773-392	Sequence 392, A
426	30	39.5	399	7	US-11-096-568A-7756	Sequence 7756, A	499	30	39.5	1826	6	US-10-330-773-499	Sequence 499, A
427	30	39.5	401	6	US-10-948-053-2	Sequence 2, Appli	500	30	39.5	1896	6	US-10-877-346-13	Sequence 13, A
428	30	39.5	406	6	US-10-821-234-1349	Sequence 1349, A							
429	30	39.5	406	6	US-10-878-568A-192	Sequence 192, A							
430	30	39.5	414	7	US-11-096-568A-28872	Sequence 28872, A							
431	30	39.5	415	7	US-11-087-099-11244	Sequence 11244, A							
432	30	39.5	431	7	US-11-096-568A-7755	Sequence 7755, A							
433	30	39.5	439	7	US-11-087-099-1909	Sequence 1909, A							
434	30	39.5	453	7	US-11-087-099-10922	Sequence 10922, A							
435	30	39.5	480	7	US-11-096-568A-7754	Sequence 7754, A							
436	30	39.5	496	7	US-11-096-568A-30934	Sequence 30934, A							
437	30	39.5	499	7	US-11-096-568A-29032	Sequence 29032, A							
438	30	39.5	509	7	US-11-055-309A-12	Sequence 12, Appl							
439	30	39.5	509	7	US-11-072-175-219	Sequence 219, A							
440	30	39.5	513	7	US-11-096-568A-28217	Sequence 28217, A							
441	30	39.5	532	6	US-10-821-234-318	Sequence 318, A							
442	30	39.5	535	7	US-11-096-568A-30933	Sequence 30933, A							
443	30	39.5	550	7	US-11-096-568A-30932	Sequence 30932, A							
444	30	39.5	555	7	US-11-096-568A-28216	Sequence 28216, A							
445	30	39.5	574	6	US-10-063-703-164	Sequence 164, A							
446	30	39.5	574	7	US-11-102-240-164	Sequence 164, A							
447	30	39.5	574	7	US-11-124-925-5	Sequence 5, Appli							
448	30	39.5	608	7	US-11-072-512-3337	Sequence 3337, A							
449	30	39.5	610	7	US-11-096-568A-14663	Sequence 14663, A							
450	30	39.5	619	7	US-10-453-372-268	Sequence 268, A							
451	30	39.5	629	6	US-11-096-568A-28295	Sequence 28295, A							
452	30	39.5	638	7	US-11-096-568A-28590	Sequence 28590, A							
453	30	39.5	647	7	US-11-096-568A-28590	Sequence 28590, A							
454	30	39.5	648	7	US-11-072-512-2032	Sequence 2032, A							
455	30	39.5	652	7	US-11-096-568A-28294	Sequence 28294, A							
456	30	39.5	661	7	US-11-096-568A-28589	Sequence 28589, A							
457	30	39.5	716	7	US-11-105-259-2	Sequence 2, Appli							
458	30	39.5	724	6	US-10-821-234-1506	Sequence 1506, A							
459	30	39.5	731	7	US-11-130-821-3	Sequence 3, Appli							
460	30	39.5	742	6	US-10-453-372-264	Sequence 264, A							
461	30	39.5	748	6	US-10-330-773-394	Sequence 394, A							
462	30	39.5	748	7	US-11-096-568A-28293	Sequence 28293, A							
463	30	39.5	753	7	US-11-096-568A-29031	Sequence 29031, A							

ALIGNMENTS

RESULT 1

US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication NO. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlögel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT FILING DATE: 2005-07-21
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760

```
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match      100.0%; Score 76; DB 7; Length 760;
Best Local Similarity 100.0%; Pred. No. 2,1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEENT 15
Db 24 IVLRPSRVHSEENT 38

RESULT 2
US-11-080-991-54
; Sequence 54, Application US/11080991
; Publication No. US2005026437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 2828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-080-991-54

Query Match      53.9%; Score 41; DB 7; Length 2828;
Best Local Similarity 35.7%; Pred. No. 98;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEENT 14
Db 2238 VMKPAKIEHKEEN 2251

RESULT 3
US-11-186-284-49
; Sequence 49, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPW01-029P2RM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 49
; LENGTH: 2828
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-49

Query Match      53.9%; Score 41; DB 7; Length 2828;
Best Local Similarity 35.7%; Pred. No. 98;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHSEEN 14
Db 2238 VMKPAKIEHKEEN 2251

RESULT 4
US-11-096-568A-21152
; Sequence 21152, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21152
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(331)
; OTHER INFORMATION: Ceres Seq. ID no. 12401933
US-11-096-568A-21152

Query Match      52.6%; Score 40; DB 7; Length 331;
Best Local Similarity 54.5%; Pred. No. 14;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNS 11
Db 259 VMLRPLQHNS 269

RESULT 5
US-11-096-568A-21150
; Sequence 21150, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21150
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(377)
; OTHER INFORMATION: Ceres Seq. ID no. 12401931
US-11-096-568A-21150

Query Match      52.6%; Score 40; DB 7; Length 377;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNS 11
```



```
Db          305 VMLRPLQIHNS 315
::||| ::|||
RESULT 6
US-11-087-099-7794
; Sequence 7794, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7794
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Daucus carota
US-11-087-099-7794
Query Match          52.6%; Score 40; DB 7; Length 532;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy          3 LRPSRVHNSSE 15
|:|:|:|:|:|
Db          56 LKPRQVHRPESNT 68

RESULT 7
US-11-096-568A-32095
; Sequence 32095, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32095
; LENGTH: 917
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc feature
; LOCATION: (1)-(917)
; OTHER INFORMATION: Ceres Seq. ID no. 15221080
US-11-096-568A-32095
Query Match          51.3%; Score 39; DB 7; Length 917;
Best Local Similarity 35.7%; Pred. No. 64;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy          1 IVLRPSRVHNSSE 14
:|:|:|:|:|
Db          47 VFLKPAKVHDDDED 60

RESULT 8
US-11-096-568A-32094
; Sequence 32094, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
```

```
; SEQ ID NO 32094
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(949)
; OTHER INFORMATION: Ceres Seq. ID no. 15221079
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-32094
Query Match          51.3%; Score 39; DB 7; Length 949;
Best Local Similarity 35.7%; Pred. No. 66;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy          1 IVLRPSRVHNSSE 14
:|:|:|:|:|
Db          79 VFLKPAKVHDDDED 92

RESULT 9
US-11-087-099-11054
; Sequence 11054, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11054
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-087-099-11054
Query Match          50.0%; Score 38; DB 7; Length 252;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy          3 LRPSRVHNSSE 12
||| ||| |||
Db          14 LRPRGVNSSE 23

RESULT 10
US-11-096-568A-28365
; Sequence 28365, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28365
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(507)
; OTHER INFORMATION: Ceres Seq. ID no. 2715595
US-11-096-568A-28365
Query Match          48.7%; Score 37; DB 7; Length 507;
Best Local Similarity 54.5%; Pred. No. 75;
```

```
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
   ||| :||| :
Db 23 LQPSHLHSRQ 33

RESULT 11
US-11-096-568A-28364
; Sequence 28364, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28364
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(632)
; OTHER INFORMATION: Ceres Seq. ID no. 2715594
US-11-096-568A-28364

Query Match 48.7%; Score 37; DB 7; Length 632;
Best Local Similarity 54.5%; Pred. No. 95;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
   ||| :||| :
Db 148 LQPSHLHSRQ 158

RESULT 12
US-11-096-568A-28363
; Sequence 28363, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28363
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(732)
; OTHER INFORMATION: Ceres Seq. ID no. 2715593
US-11-096-568A-28363

Query Match 48.7%; Score 37; DB 7; Length 732;
Best Local Similarity 54.5%; Pred. No. 11e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSE 13
   ||| :||| :
Db 248 LQPSHLHSRQ 258

RESULT 13
US-11-087-099-2391
; Sequence 2391, Application US/11087099
; Publication No. US20060041961A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2391
; LENGTH: 2479
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-11-087-099-2391

Query Match 48.7%; Score 37; DB 7; Length 2479;
Best Local Similarity 61.5%; Pred. No. 4.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLPRSRVHNSSE 14
   |||| :||| :
Db 245 VLPRPSRLNSDN 257

RESULT 14
US-10-935-494-3
; Sequence 3, Application US/10935494
; Publication No. US20060024329A1
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200.1
; CURRENT APPLICATION NUMBER: US/10/935,494
; CURRENT FILING DATE: 2004-09-07
; PRIOR APPLICATION NUMBER: US/09/596,479
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/213,053
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 9608242
; PRIOR FILING DATE: 1996-06-27
; PRIOR APPLICATION NUMBER: PCT/FR97/01115
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Canine herpesvirus
US-10-935-494-3

Query Match 47.4%; Score 36; DB 6; Length 420;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 14
   |||| :||| :
Db 80 ITNHPSQIINNEEN 93

RESULT 15
US-11-124-367A-473
; Sequence 473, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
```

```
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 473
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-367A-473

Query Match 47.4%; Score 36; DB 7; Length 821;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 PSRVHSEENT 15
Db 503 PLRIHEKEET 513

RESULT 16
US-10-935-494-31
; Sequence 31, Application US/10935494
; Publication No. US20060024329A1
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BAUDU, Philippe
; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
; TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200.1
; CURRENT APPLICATION NUMBER: US/10/935,494
; CURRENT FILING DATE: 2004-09-07
; PRIOR APPLICATION NUMBER: US/09/596,479
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/213,053
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 9608242
; PRIOR FILING DATE: 1996-06-27
; PRIOR APPLICATION NUMBER: PCF/FR97/01115
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 31
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Canine herpesvirus
US-10-935-494-31

Query Match 47.4%; Score 36; DB 6; Length 1394;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IIVLRPSRVHSEENT 14
Db 80 ITNHPSQIINSEEN 93

RESULT 17
US-11-087-099-835
; Sequence 835, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21153450(B) EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 835
; LENGTH: 205

; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 473
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-367A-473

Query Match 46.1%; Score 35; DB 7; Length 205;
Best Local Similarity 63.6%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHSEENT 14
Db 4 RPSRIECSEPN 14

RESULT 18
US-11-234-786-590
; Sequence 590, Application US/11234786
; Publication No. US20060024301A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Sreiky, Yasir A.
; TITLE OF INVENTION: PROSTATE-SPECIFIC POLYPEPTIDES AND FUSION
; TITLE OF INVENTION: POLYPEPTIDES THEREOF
; FILE REFERENCE: 210121.427C31
; CURRENT APPLICATION NUMBER: US/11/234,786
; CURRENT FILING DATE: 2005-09-23
; PRIOR APPLICATION NUMBER: US 09/568,857
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US 09/536,857
; PRIOR FILING DATE: 2000-05-27
; PRIOR APPLICATION NUMBER: US 09/483,672
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 09/439,313
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/352,616
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: US 09/288,946
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/232,149
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: US 09/159,812
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: US 09/115,453
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: US 09/030,607
; PRIOR FILING DATE: 1998-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 701
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 590
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-234-786-590

Query Match 46.1%; Score 35; DB 7; Length 347;
Best Local Similarity 41.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RPSRVHSEENT 15
```

```
Db 320 RSNRIHQKQNT 331
; :|:|:|:|:|
; LOCATION: (325)..(325)
; OTHER INFORMATION: Xaa is any aa, unknown or other
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (538)..(538)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-11818
Query Match 46.1%; Score 35; DB 7; Length 555;
Best Local Similarity 61.5%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 13
| | | | |
Db 439 IVLRGCRVGSDE 451
| | | | |

RESULT 21
US-11-087-099-829
; Sequence 829, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 829
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Brassica rapa
US-11-087-099-829
Query Match 46.1%; Score 35; DB 7; Length 838;
Best Local Similarity 61.5%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VLRPSRVHNSSE 14
| | | | |
Db 479 VVPSRRHNSRE 491
| | | | |

RESULT 22
US-11-087-099-1383
; Sequence 1383, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1383
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-11-087-099-1383
Query Match 44.7%; Score 34; DB 7; Length 181;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 14
| | | | |
Db 7 IYVRPLEVTDAAEN 20
| | | | |

RESULT 23
US-11-096-568A-5328
; Sequence 5328, Application US/11096568A
; Publication No. US20060048240A1
```

```
Db 320 RSNRIHQKQNT 331
; :|:|:|:|:|
; LOCATION: (325)..(325)
; OTHER INFORMATION: Xaa is any aa, unknown or other
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (538)..(538)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-11818
Query Match 46.1%; Score 35; DB 7; Length 542;
Best Local Similarity 61.5%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSE 13
| | | | |
Db 426 IVLRGCRVGSDE 438
| | | | |

RESULT 20
US-11-096-568A-11818
; Sequence 11818, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 11818
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(555)
; OTHER INFORMATION: Ceres Seq. ID no. 15220615
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (324)..(324)
; OTHER INFORMATION: Xaa is any aa, unknown or other
; FEATURE:
; NAME/KEY: misc_feature
```

```
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5328
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(363)
; OTHER INFORMATION: Ceres Seq. ID no. 14308169
US-11-096-568A-5328

Query Match          44.7%; Score 34; DB 7; Length 363;
Best Local Similarity 46.2%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 LRPSRVHNSSENT 15
      ||||| : ||| :
Db      41 LRPSMIKNKERS 53

RESULT 24
US-11-124-368A-209
; Sequence 209, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR FILING DATE: 2005-05-07
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-209

Query Match          44.7%; Score 34; DB 7; Length 377;
Best Local Similarity 41.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      1 IVLRPSRVHNS 12
      :||| : ||| :
Db      95 VLIRLSRLHNOQ 106

RESULT 25
US-11-124-368A-210
; Sequence 210, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
```

```
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-210

Query Match          44.7%; Score 34; DB 7; Length 377;
Best Local Similarity 41.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      1 IVLRPSRVHNS 12
      :||| : ||| :
Db      95 VLIRLSRLHNOQ 106

RESULT 26
US-11-087-099-4550
; Sequence 4550, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4550
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Corynebacterium efficiens YS-314
US-11-087-099-4550

Query Match          44.7%; Score 34; DB 7; Length 486;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      3 LRPSRVHNSSENT 15
      ||||| : ||| :
Db      408 LRPLEHNGEIST 420

RESULT 27
US-11-067-231-153
; Sequence 153, Application US/11067231
; Publication No. US20050272063A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Katagiri, Toyomasa
; APPLICANT: Fukukawa, Chikako
; TITLE OF INVENTION: METHOD FOR TREATING SYNOVIAL SARCOMA
; FILE REFERENCE: 1254-0272PUS1
; CURRENT APPLICATION NUMBER: US/11/067,231
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/407,506
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/486,195
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/JP03/10591
; PRIOR FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: PCT/JP2004/002144
; PRIOR FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US 60/598,834
; PRIOR FILING DATE: 2004-08-05
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 153
```

```

; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-231-153

Query Match 44.7%; Score 34; DB 7; Length 581;
Best Local Similarity 38.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLRPSRVHNSSEN 14
: ||| |:::|
Db 168 LFRPQPHSAQEH 180

RESULT 28
US-11-054-281-22
; Sequence 22, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-281-22

Query Match 44.7%; Score 34; DB 7; Length 581;
Best Local Similarity 38.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLRPSRVHNSSEN 14
: ||| |:::|
Db 168 LFRPQPHSAQEH 180

RESULT 29
US-11-054-281-91
; Sequence 91, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-054-281-22

Query Match 44.7%; Score 34; DB 7; Length 581;
Best Local Similarity 38.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLRPSRVHNSSEN 14
: ||| |:::|
Db 168 LFRPQPHSAQEH 180

RESULT 30
US-10-793-626-3328
; Sequence 3328, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3328
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3328

Query Match 44.7%; Score 34; DB 6; Length 588;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LPRSRVHNS 12
: |||||
Db 183 LPRSRVHNS 192

RESULT 31
US-11-087-099-3671
; Sequence 3671, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3671
; LENGTH: 593
; TYPE: PRT

```

```
; ORGANISM: Schizosaccharomyces pombe
US-11-087-099-3671

Query Match      44.7%; Score 34; DB 7; Length 593;
Best Local Similarity 53.3%; Pred. No. 2.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 3 LRP-----SRVHNSSE 13
Db 53 LRPDNEFAEVHNSD 67

RESULT 32
US-11-087-099-9040
; Sequence 9040, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abed, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9040
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-11-087-099-9040

Query Match      44.7%; Score 34; DB 7; Length 594;
Best Local Similarity 53.3%; Pred. No. 2.9e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

Qy 3 LRP-----SRVHNSSE 13
Db 53 LRPDNEFAEVHNSD 67

RESULT 33
US-10-055-877-157
; Sequence 157, Application US/10055877
; Publication No. US20050289241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Keku, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Caaman, Stacie
; APPLICANT: Boldog, Ferenc
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby

; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 157
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-157

Query Match      44.7%; Score 34; DB 6; Length 757;
Best Local Similarity 33.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSENT 15
Db 276 LVNRKKNHNLNENT 290

RESULT 34
US-10-821-234-1592
; Sequence 1592, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1592
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1592

Query Match      44.7%; Score 34; DB 6; Length 782;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 SRVHNSSENT 15
Db 256 ARVHVSSEGT 265
```

RESULT 35
US-11-177-506-38
; Sequence 38, Application US/11177506
; Publication No. US20060029956A1
; GENERAL INFORMATION:
; APPLICANT: Beyer, Wayne F.
; APPLICANT: Venetta, Thomas M.
; APPLICANT: Groelke, John W.
; APPLICANT: Blaesus, Rainer H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF OVARIAN DISEASE
; FILE REFERENCE: 46143/294851
; CURRENT APPLICATION NUMBER: US/11/177.506
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: 60/586,856
; PRIOR FILING DATE: 2004-07-09
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-177-506-38

Query Match 44.7%; Score 34; DB 7; Length 782;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 SRVNSENT 15
Db 256 ARVHSEGT 265

RESULT 36
US-10-509-422-5
; Sequence 5, Application US/10509422
; Publication No. US20050244825A1
; GENERAL INFORMATION:
; APPLICANT: Liou, Simon
; TITLE OF INVENTION: Human BMP2 Inducible Kinases
; FILE REFERENCE: 004974.01015
; CURRENT APPLICATION NUMBER: US/10/509,422
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/EP03/080825
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/367,512
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/406,936
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-509-422-5

Query Match 44.7%; Score 34; DB 6; Length 1311;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
Db 461 RPSRFHN 467

RESULT 37
US-11-179-624-3
; Sequence 3, Application US/11179624
; Publication No. US20060040372A1
; GENERAL INFORMATION:
; APPLICANT: Sakaguchi, Nobuo
; APPLICANT: Kuwahara, Kazuhiko

; TITLE OF INVENTION: GANP Protein
; FILE REFERENCE: 050208-0018
; CURRENT APPLICATION NUMBER: US/11/179,624
; CURRENT FILING DATE: 2005-07-13
; PRIOR APPLICATION NUMBER: PCT/JP99/04634
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/914,272
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 10/747,133
; PRIOR FILING DATE: 2003-12-30
; PRIOR APPLICATION NUMBER: 47035/1999
; PRIOR FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1980
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-179-624-3

Query Match 44.7%; Score 34; DB 7; Length 1980;
Best Local Similarity 56.2%; Pred. No. 1.1e+03;
Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 2 VLPRSRVH--NSENT 15
Db 384 VLAPSRIPGVNKEET 399

RESULT 38
US-10-884-730-115
; Sequence 115, Application US/10884730
; Publication No. US20050250161A1
; GENERAL INFORMATION:
; APPLICANT: Suciu-Foca, Nicole
; APPLICANT: Liou, Zhuoro
; APPLICANT: Chang, Chih-Chao
; APPLICANT: Cortesini, Raffaello
; TITLE OF INVENTION: Generation of Antigen-Specific T Suppressor Cells For Treatment
; TITLE OF INVENTION: Rejection
; FILE REFERENCE: 0575/58332-B
; CURRENT APPLICATION NUMBER: US/10/884,730
; CURRENT FILING DATE: 2004-07-02
; PRIOR APPLICATION NUMBER: US/09/746,311
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/16594
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 382
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 115
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Human HLA-DRB
US-10-884-730-115

Query Match 43.4%; Score 33; DB 6; Length 89;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 RVHNSSE 13
Db 25 RVHNSSE 31

RESULT 39
US-11-096-568A-27299
; Sequence 27299, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2

;; CURRENT APPLICATION NUMBER: US/11/096,568A
;; CURRENT FILING DATE: 2005-04-01
;; NUMBER OF SEQ ID NOS: 34471
;; SEQ ID NO 27299
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Arabidopsis thaliana
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(173)
;; OTHER INFORMATION: Ceres Seq. ID no. 13610690
US-11-096-568A-27299

Query Match 43.4%; Score 33; DB 7; Length 173;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LRPSRVHNSSENT 15
||| ||| : :
Db 10 LRPSLVHDFLOQT 22

RESULT 40
US-11-096-568A-328
; Sequence 328, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 328
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(180)
; OTHER INFORMATION: Ceres Seq. ID no. 15179512
US-11-096-568A-328

Query Match 43.4%; Score 33; DB 7; Length 180;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
||| |||
Db 171 RPLRVHN 177

RESULT 41
US-11-087-099-5689
; Sequence 5689, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5689
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(184)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-5689

Query Match 43.4%; Score 33; DB 7; Length 184;
Best Local Similarity 33.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLRPSRVHNSSENT 15
| : : : : :
Db 47 IITKPKVHSTTRAT 61

RESULT 42
US-11-096-568A-327
; Sequence 327, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 327
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(231)
; OTHER INFORMATION: Ceres Seq. ID no. 15179511
US-11-096-568A-327

Query Match 43.4%; Score 33; DB 7; Length 231;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
||| |||
Db 222 RPLRVHN 228

RESULT 43
US-11-096-568A-329
; Sequence 329, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 329
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(231)
; OTHER INFORMATION: Ceres Seq. ID no. 16625618
US-11-096-568A-329

Query Match 43.4%; Score 33; DB 7; Length 231;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
||| |||
Db 222 RPLRVHN 228

RESULT 44

US-10-793-626-1894
; Sequence 1894, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn, Ver. 2.1
; SEQ ID NO 1894
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1894

Query Match 43.4%; Score 33; DB 6; Length 241;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSE 13
: : : : :
Db 213 ILFEPKIDNTEE 225

RESULT 45
US-10-884-730-114
; Sequence 114, Application US/10884730
; Publication No. US20050250161A1
; GENERAL INFORMATION:
; APPLICANT: Suciu-Foca, Nicole
; APPLICANT: Liu, Zhuoro
; APPLICANT: Chang, Chih-Chao
; APPLICANT: Cortesini, Raffaello
; TITLE OF INVENTION: Generation of Antigen-Specific T Suppressor Cells For Treatment of
; TITLE OF INVENTION: Rejection
; FILE REFERENCE: 0575/58332-B
; CURRENT APPLICATION NUMBER: US/10/884,730
; CURRENT FILING DATE: 2004-07-02
; PRIOR APPLICATION NUMBER: US/09/746,311
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 382
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 114
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Human HLA-DRB
US-10-884-730-114

Query Match 43.4%; Score 33; DB 6; Length 266;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 RVHNSSE 13
: : : : :
Db 59 RVHNSSE 65

RESULT 46
US-11-087-099-287
; Sequence 287, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement

; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 287
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-287

Query Match 43.4%; Score 33; DB 7; Length 287;
Best Local Similarity 46.2%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNSSE 13
: : : : :
Db 255 VEMKRRVHREE 267

RESULT 47
US-11-096-568A-19351
; Sequence 19351, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19351
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: Ceres Seq. ID no. 12372810
US-11-096-568A-19351

Query Match 43.4%; Score 33; DB 7; Length 289;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
: : : : :
Db 106 RPSRVHN 112

RESULT 48
US-11-096-568A-19350
; Sequence 19350, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19350
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(295)
; OTHER INFORMATION: Ceres Seq. ID no. 12372809
US-11-096-568A-19350

Query Match 43.4%; Score 33; DB 7; Length 295;

Best Local Similarity 85.7%; Pred. No. 2e+02; Mismatches 0; Indels 1; Gaps 0;
Matches 6; Conservative

Qy 4 RPSRVHN 10
Db 112 RPSRVHN 118

RESULT 49

US-11-024-959-267
; Sequence 267, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; PRIOR FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 267
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-267

Query Match 43.4%; Score 33; DB 7; Length 313;
Best Local Similarity 38.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNS 13
Db 41 VALKKTRLHDEE 53

RESULT 50

US-11-096-568A-19349
; Sequence 19349, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19349
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(315)
; OTHER INFORMATION: Ceres Seq. ID no. 12372808
US-11-096-568A-19349

Query Match 43.4%; Score 33; DB 7; Length 315;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHN 10
Db 132 RPSRVHN 138

RESULT 51

US-11-096-568A-12602
; Sequence 12602, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12602
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(339)
; OTHER INFORMATION: Ceres Seq. ID no. 14302455
US-11-096-568A-12602

Query Match 43.4%; Score 33; DB 7; Length 339;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RPSRVHNS 11
Db 10 RPSRVHNS 17

RESULT 52

US-10-467-657-5522
; Sequence 5522, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5522
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5522

Query Match 43.4%; Score 33; DB 6; Length 387;
Best Local Similarity 35.7%; Pred. No. 2.7e+02;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVLPRSRVHNS 14
Db 1 VVVRQSIHHWTDN 14

RESULT 53

US-11-069-642-92
; Sequence 92, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.

APPLICANT: KINSELLA, TODD M.
APPLICANT: BENNETT, MARK K.
TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
FILE REFERENCE: RIGL-022CIP3
CURRENT APPLICATION NUMBER: US/11/069,642
CURRENT FILING DATE: 2005-02-28
PRIOR APPLICATION NUMBER: 10/232,758
PRIOR FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 10/422,536
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 09/800,770
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/187,130
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 168
SOFTWARE: PatentIn version 3.2
SEQ ID NO 92
LENGTH: 437
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: DnaB intein cyclization scaffold with GFP
US-11-069-642-92

Query Match 43.4%; Score 33; DB 7; Length 437;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VHNSEEN 14
Db 50 VHNSEED 56

RESULT 54
US-11-069-642-47
Sequence 47, Application US/11069642
Publication No. US20050260626A1
GENERAL INFORMATION:
APPLICANT: LORENS, JAMES B.
APPLICANT: PRAY, TODD R.
APPLICANT: KINSELLA, TODD M.
APPLICANT: BENNETT, MARK K.
TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
FILE REFERENCE: RIGL-022CIP3
CURRENT APPLICATION NUMBER: US/11/069,642
CURRENT FILING DATE: 2005-02-28
PRIOR APPLICATION NUMBER: 10/232,758
PRIOR FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 10/422,536
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 09/800,770
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/187,130
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 168
SOFTWARE: PatentIn version 3.2
SEQ ID NO 47
LENGTH: 438
TYPE: PRT
ORGANISM: Synecocystis PCC6803
US-11-069-642-47

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VHNSEEN 14
Db 50 VHNSEED 56

RESULT 55
US-11-069-642-49
Sequence 49, Application US/11069642
Publication No. US20050260626A1
GENERAL INFORMATION:
APPLICANT: LORENS, JAMES B.
APPLICANT: PRAY, TODD R.
APPLICANT: KINSELLA, TODD M.
APPLICANT: BENNETT, MARK K.
TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
FILE REFERENCE: RIGL-022CIP3
CURRENT APPLICATION NUMBER: US/11/069,642
CURRENT FILING DATE: 2005-02-28
PRIOR APPLICATION NUMBER: 10/232,758
PRIOR FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 10/422,536
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 09/800,770
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/187,130
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 168
SOFTWARE: PatentIn version 3.2
SEQ ID NO 49
LENGTH: 438
TYPE: PRT
ORGANISM: Synecocystis PCC6803
US-11-069-642-49

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VHNSEEN 14
Db 50 VHNSEED 56

RESULT 56
US-11-069-642-51
Sequence 51, Application US/11069642
Publication No. US20050260626A1
GENERAL INFORMATION:
APPLICANT: LORENS, JAMES B.
APPLICANT: PRAY, TODD R.
APPLICANT: KINSELLA, TODD M.
APPLICANT: BENNETT, MARK K.
TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
FILE REFERENCE: RIGL-022CIP3
CURRENT APPLICATION NUMBER: US/11/069,642
CURRENT FILING DATE: 2005-02-28
PRIOR APPLICATION NUMBER: 10/232,758
PRIOR FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 10/422,536
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 09/800,770
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: 60/187,130
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 168
SOFTWARE: PatentIn version 3.2
SEQ ID NO 51
LENGTH: 438
TYPE: PRT
ORGANISM: Synecocystis PCC6803
US-11-069-642-51

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VHNSEEN 14
|||||:
Db 50 VHNSEED 56

RESULT 57

US-11-069-642-53

; Sequence 53, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Synechocystis PCC6803
US-11-069-642-53

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VHNSEEN 14
|||||:
Db 50 VHNSEED 56

RESULT 58

US-11-069-642-55
; Sequence 55, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 55
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Synechocystis PCC6803
US-11-069-642-55

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VHNSEEN 14
|||||:
Db 50 VHNSEED 56

RESULT 59

US-11-069-642-57
; Sequence 57, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Synechocystis PCC6803
US-11-069-642-57

Query Match 43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VHNSEEN 14
|||||:
Db 50 VHNSEED 56

RESULT 60

US-11-069-642-59
; Sequence 59, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; APPLICANT: BENNETT, MARK K.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; TITLE OF INVENTION: INHIBITING PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
US-11-069-642-59

```
; SEQ ID NO 59
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Synchocystis PCC6803
US-11-069-642-59

Query Match      43.4%; Score 33; DB 7; Length 438;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      8 VHNSEN 14
Db      50 VHNSEED 56
```

Search completed: March 11, 2006, 12:32:55
Job time : 17.8158 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:11:52 ; Search time 121.447 Seconds
(without alignments)
47.032 Million cell updates/sec

Title: US-10-774-242A-6
Perfect score: 62
Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	13	ADR47262	Adr47262 Human alp
2	62	100.0	723	AEb94227	Aeb94227 Human sol
3	62	100.0	734	AEb94218	Aeb94218 Human sol
4	62	100.0	750	AEb94161	Aeb94161 Human sol
5	62	100.0	759	AAW31963	Aaw31963 Human fib
6	62	100.0	760	AAW27438	Aaw27438 Human fib
7	62	100.0	760	ABR47452	AbR47452 Breast ca
8	62	100.0	760	ADN95552	Adn95552 Human BEC
9	62	100.0	760	ADQ21351	Adq21351 Human sof
10	62	100.0	760	ADW14775	Adw14775 Tumor-ass
11	62	100.0	760	AEb94159	Aeb94159 Human wil
12	55	88.7	761	AEb94163	Aeb94163 Mouse wil
13	49	79.0	562	ADX90113	Adx90113 Plant ful
14	43	69.4	223	AAW20957	Aaw20957 H. pylori
15	43	69.4	296	AAW20427	Aaw20427 H. pylori
16	43	69.4	345	ABUS1130	Abu51130 Helicobac
17	43	69.4	346	ABUS2523	Abu52523 Helicobac
18	43	69.4	706	ABR531675	AbR531675 Protein s
19	43	69.4	706	ADB96560	Adb96560 Yeast his
20	43	69.4	706	ADK64084	Adk64084 Disease c
21	42	67.7	17	ADL70610	Adl70610 Human thr
22	42	67.7	18	ADL70611	Adl70611 Human thr
23	42	67.7	52	AAO04456	Aao04456 Human pol
24	42	67.7	98	ADL70643	Adl70643 Human thr

98	40	64.5	962	9	ADW11776	Adw11776 Human TAP	171	38	61.3	536	6	ABU40906	Abu40906 Protein e
99	40	64.5	990	8	AD061721	Ad061721 Transcrip	172	38	61.3	542	7	AD516661	Ad516661 B. hensel
100	40	64.5	990	8	AD080705	Ad080705 Arabidops	173	38	61.3	569	7	Adf06692	Adf06692 Bacterial
101	40	64.5	1075	3	AG52001	Ag52001 Arabidops	174	38	61.3	609	6	ABR52773	ABR52773 Protein s
102	40	64.5	1077	3	AG52000	Ag52000 Arabidops	175	38	61.3	609	6	ADK62122	ADK62122 Disease t
103	40	64.5	1164	3	AG51999	Ag51999 Arabidops	176	38	61.3	609	8	ADN18927	ADN18927 Bacterial
104	40	64.5	1279	4	AB65886	Ab65886 Drosophil	177	38	61.3	617	7	ABO74672	ABO74672 Pseudomn
105	40	64.5	2285	2	AM98149	Am98149 Bacillus	178	38	61.3	711	8	ABM84518	ABM84518 Human dia
106	39	62.9	50	4	ABB40235	Abb40235 Peptide #	179	38	61.3	715	7	ADC37558	ADC37558 Human nuc
107	39	62.9	50	4	AM33908	Am33908 Peptide #	180	38	61.3	762	4	ABU35926	Abu35926 Helicobac
108	39	62.9	50	4	ABB24665	Abb24665 Peptide #	181	38	61.3	762	4	ABU30830	Abu30830 Protein e
109	39	62.9	50	4	AM73721	Am73721 Human bon	182	38	61.3	779	9	ADY70376	Ady70376 Human bet
110	39	62.9	50	4	AM61021	Am61021 Human bra	183	38	61.3	790	8	ADN21150	ADN21150 Bacterial
111	39	62.9	50	4	ABG55462	Abg55462 Human liv	184	38	61.3	797	8	ABM84517	ABM84517 Human dia
112	39	62.9	50	5	ABG43599	Abg43599 Human pep	185	38	61.3	799	8	ADQ90745	ADQ90745 Human ace
113	39	62.9	161	8	ADS17865	AdS17865 Human iKB	186	38	61.3	802	8	ABM84516	ABM84516 Human dia
114	39	62.9	230	3	AG520298	Ag520298 Arabidops	187	38	61.3	843	7	ADN90890	ADN90890 Human hep
115	39	62.9	230	3	AG51753	Ag51753 Arabidops	188	38	61.3	843	7	ADN03920	ADN03920 Antipspor
116	39	62.9	233	4	AM35049	Am35049 Enterococ	189	38	61.3	843	9	ADZ87302	ADZ87302 Human com
117	39	62.9	233	4	ABU29360	Abu29360 Protein e	190	38	61.3	910	7	ABR62791	ABR62791 MRSA 2-ke
118	39	62.9	234	3	AG20297	Ag20297 Arabidops	191	38	61.3	932	2	AAW97701	AAW97701 Staphyloc
119	39	62.9	234	3	AG51752	Ag51752 Arabidops	192	38	61.3	932	4	AAU37022	AAU37022 Staphyloc
120	39	62.9	268	3	AG51751	Ag51751 Arabidops	193	38	61.3	932	6	ABM73033	ABM73033 Staphyloc
121	39	62.9	269	3	AG20296	Ag20296 Arabidops	194	38	61.3	1320	9	ADY85279	ADY85279 Rat homol
122	39	62.9	291	5	AB55586	Ab55586 Lactococc	195	38	61.3	1381	4	ABB57920	ABB57920 Drosophil
123	39	62.9	368	8	ADS17863	AdS17863 Human iKB	196	38	61.3	2346	8	ADQ90744	ADQ90744 Human ace
124	39	62.9	384	4	AM40982	Am40982 Human pol	197	38	61.3	16368	6	ABM67171	ABM67171 Photorhab
125	39	62.9	405	2	AAW55681	AAW55681 H. pylori	198	37.5	60.5	279	9	AEA22098	Aea22098 Campyloba
126	39	62.9	405	2	AAW98502	AAW98502 H. pylori	199	37	59.7	58	6	ABU24667	Abu24667 Protein e
127	39	62.9	411	7	ADC36536	DNA encod	200	37	59.7	72	3	ABU24667	Abu24667 Protein e
128	39	62.9	411	8	AD161942	Ad161942 Necrosis	201	37	59.7	99	8	ADU82602	ADU82602 A. thalia
129	39	62.9	411	8	AD161939	Ad161939 Necrosis	202	37	59.7	130	5	ABB47414	ABB47414 Listeria
130	39	62.9	411	8	AD129004	Ad129004 Human NEM	203	37	59.7	159	5	ABP04200	ABP04200 Human ORF
131	39	62.9	411	8	AD129001	Ad129001 Human NEM	204	37	59.7	173	5	ABP26684	ABP26684 Streptoco
132	39	62.9	412	9	AD200650	Ad200650 Mouse NEM	205	37	59.7	185	8	ADX68131	ADX68131 Plant ful
133	39	62.9	416	2	AAV27430	AAV27430 Human RIP	206	37	59.7	219	8	ADY09931	ADY09931 Plant ful
134	39	62.9	419	6	AB82782	AB82782 Human NEM	207	37	59.7	230	8	ADK16871	ADK16871 Nanoarcha
135	39	62.9	419	6	AB017485	AB017485 Human NEM	208	37	59.7	231	8	ADN20564	ADN20564 Bacterial
136	39	62.9	419	8	ADK71963	ADK71963 Human I K	209	37	59.7	231	8	ADY11734	ADY11734 Plant ful
137	39	62.9	419	8	AD88168	AD88168 Human pro	210	37	59.7	251	8	ADX94404	ADX94404 Plant ful
138	39	62.9	419	8	ADY04410	ADY04410 Human NEM	211	37	59.7	255	4	AAK61796	AAK61796 S. cerevis
139	39	62.9	420	8	AD200656	Ad200656 Human NEM	212	37	59.7	255	7	ADT87153	ADT87153 Yeast Str
140	39	62.9	425	4	AB60498	AB60498 Human cel	213	37	59.7	255	8	ADY10830	ADY10830 Plant ful
141	39	62.9	425	4	AAK39196	AAK39196 Human pol	214	37	59.7	262	3	AAK12782	AAK12782 Arabidops
142	39	62.9	425	7	ADM05696	ADM05696 Human pro	215	37	59.7	262	3	AAK53105	AAK53105 Arabidops
143	39	62.9	438	2	AAV35353	AAV35353 Amino aci	216	37	59.7	285	3	AAK53104	AAK53104 Arabidops
144	39	62.9	575	8	ADS29966	AdS29966 Bacterial	217	37	59.7	285	3	AAK12781	AAK12781 Arabidops
145	39	62.9	678	5	ABP41521	ABP41521 Human ova	218	37	59.7	294	3	AAK53103	AAK53103 Arabidops
146	39	62.9	790	6	ABU24454	Abu24454 Protein e	219	37	59.7	294	6	ADA15531	ADA15531 A. thalia
147	39	62.9	797	4	AB60378	AB60378 Drosophil	220	37	59.7	294	7	ADB31893	ADB31893 Plant (A.
148	39	62.9	867	6	ABU41082	Abu41082 Protein e	221	37	59.7	294	8	ADO02057	ADO02057 Thalecres
149	39	62.9	870	7	ADF04483	Adf04483 Bacterial	222	37	59.7	301	3	AAV77961	AAV77961 A. thalia
150	39	62.9	3225	7	ADJ68448	AdJ68448 Human hea	223	37	59.7	301	8	ADU92236	ADU92236 Thale cre
151	39	62.9	3259	7	AD566037	Ad566037 Human Pro	224	37	59.7	308	9	ADM17310	ADM17310 E. grandis
152	39	62.9	3259	7	AD566033	Ad566033 Human Pro	225	37	59.7	308	9	ADM17310	ADM17310 E. grandis
153	38	61.3	20	6	AB017492	AB017492 Human NEM	226	37	59.7	310	8	ADX74157	ADX74157 Thale cre
154	38	61.3	89	5	ABU50736	Abu50736 Helicobac	227	37	59.7	311	8	ADN72157	ADN72157 Thale cre
155	38	61.3	141	5	ABP03532	ABP03532 Human ORF	228	37	59.7	311	8	AEA29342	Aea29342 Brassica
156	38	61.3	152	5	ABU51474	Abu51474 Helicobac	229	37	59.7	348	8	ADT55741	ADT55741 Plant pol
157	38	61.3	165	5	ABU51727	Abu51727 Helicobac	230	37	59.7	365	8	ADP04797	ADP04797 Sea squir
158	38	61.3	212	5	ABU52128	Abu52128 Helicobac	231	37	59.7	381	8	ADK74157	ADK74157 Plant ful
159	38	61.3	215	5	ABU52164	Abu52164 Helicobac	232	37	59.7	390	8	ADK16767	ADK16767 Nanoarcha
160	38	61.3	219	5	ABU51428	Abu51428 Helicobac	233	37	59.7	416	6	ABU24301	ABU24301 Protein e
161	38	61.3	221	5	ABU51528	Abu51528 Helicobac	234	37	59.7	421	8	ADO80496	ADO80496 Clostridi
162	38	61.3	222	5	ABU51344	Abu51344 Helicobac	235	37	59.7	447	6	ABM68292	ABM68292 Photorhab
163	38	61.3	364	5	AAE25291	AAE25291 Human nuc	236	37	59.7	520	4	ABG18077	ABG18077 Novel hum
164	38	61.3	368	5	ABP74073	ABP74073 Candida a	237	37	59.7	530	3	AA805942	AA805942 Protein d
165	38	61.3	430	8	ADO80484	ADO80484 Thermotog	238	37	59.7	588	8	ADT56784	ADT56784 Plant pol
166	38	61.3	430	8	ADN20325	ADN20325 Bacterial	239	37	59.7	682	8	ADX92461	ADX92461 Plant ful
167	38	61.3	446	3	AA803441	AA803441 Porcine c	240	37	59.7	762	8	ADY10146	ADY10146 Plant ful
168	38	61.3	448	2	AAV37201	AAV37201 Chlamydia	241	37	59.7	791	8	ADY22723	ADY22723 Plant ful
169	38	61.3	448	6	ABU27207	Abu27207 Protein e	242	37	59.7	911	5	ABP29744	ABP29744 Streptoco
170	38	61.3	475	8	ADK165599	ADK165599 Nanoarcha	243	37	59.7				

244	37	59.7	911	5	ABP26657	Abp26657 Streptococ	317	36	58.1	560	8	ADO19412	Ado19412 Human PRO
245	37	59.7	1279	2	AAW98761	Aaw98761 H. pylori	318	36	58.1	560	8	ADR97411	Adr97411 Human CDC
246	37	59.7	1847	3	AAV51631	Aav51631 M. jannas	319	36	58.1	560	8	ADW09038	Adw09038 Full leng
247	37	59.7	1847	3	AAV52002	Aav52002 M. jannas	320	36	58.1	560	9	ADX06784	Adx06784 Cyclin-de
248	37	59.7	1847	3	ADN18422	Adn18422 Bacterial	321	36	58.1	560	9	ADZ64349	Adz64349 Human can
249	37	59.7	2954	2	AAV01632	Aav01632 Amino aci	322	36	58.1	560	9	AEA15082	Aea15082 Human pol
250	37	59.7	3457	2	AAW84560	Aaw84560 Polyprote	323	36	58.1	575	8	ADS08188	Ads08188 Staphyloc
251	36	58.1	71	8	ABO53980	Abos3980 Human gen	324	36	58.1	582	6	ABU24101	Abu24101 Protein e
252	36	58.1	104	5	ABP00700	Abp00700 Human ORF	325	36	58.1	582	9	AEb73237	Aeb73237 Segment o
253	36	58.1	129	6	ABW71381	Abw71381 Staphyloc	326	36	58.1	582	9	AEb73315	Aeb73315 Segment o
254	36	58.1	147	3	ABE53579	Abe53579 Human col	327	36	58.1	591	8	ADX73256	Adx73256 Plant ful
255	36	58.1	148	7	ADF07371	Adf07371 Bacterial	328	36	58.1	613	2	AAW22051	Aaw22051 DNA polym
256	36	58.1	152	4	ABB26074	Abb26074 Drosophil	329	36	58.1	613	2	AAW73097	Aaw73097 Pfu DNA p
257	36	58.1	153	6	ADA34916	Ada34916 Acinetoba	330	36	58.1	613	3	AAV52020	Aav52020 F. furios
258	36	58.1	155	6	ABU24169	Abu24169 Protein e	331	36	58.1	622	4	ABB62816	Abb62816 Drosophil
259	36	58.1	176	4	ABE70814	AbE70814 Drosophil	332	36	58.1	622	9	ABE44574	Abe44574 Sleep dis
260	36	58.1	182	8	ADW89454	Adw89454 Streptoco	333	36	58.1	649	6	ABU43940	Abu43940 Protein e
261	36	58.1	182	8	ADW80707	Adw80707 Streptoco	334	36	58.1	685	8	ADN19873	Adn19873 Bacterial
262	36	58.1	210	5	ABP30191	Abp30191 Streptoco	335	36	58.1	706	8	ADO16836	Ado16836 Amylase e
263	36	58.1	210	5	ABP30191	Abp30191 Streptoco	336	36	58.1	761	8	ADN17476	Adn17476 Bacterial
264	36	58.1	217	4	ABG26744	Abg26744 Novel hum	337	36	58.1	774	3	AAW30779	Aaw30779 Arabidops
265	36	58.1	217	5	ABG28023	Abg28023 Streptoco	338	36	58.1	776	3	AAW30778	Aaw30778 Arabidops
266	36	58.1	227	3	AAV87950	Aav87950 Human cdr	339	36	58.1	780	3	AAW30777	Aaw30777 Arabidops
267	36	58.1	245	7	ADB67134	Adb67134 Tropomyos	340	36	58.1	784	8	ADM98963	Adm98963 Diterpene
268	36	58.1	245	7	ADK65795	Adk65795 Angiogene	341	36	58.1	784	8	ADM98964	Adm98964 Diterpene
269	36	58.1	279	8	ADL99405	Adl99405 Nanostruc	342	36	58.1	790	7	ADD30777	Add30777 Plant yie
270	36	58.1	281	7	ADP67132	Adp67132 Tropomyos	343	36	58.1	790	8	ADI43921	Adi43921 Plant tra
271	36	58.1	281	8	ADL99403	Adl99403 Nanostruc	344	36	58.1	856	4	ABB59118	Abb59118 Drosophil
272	36	58.1	306	4	AAW74061	Aaw74061 Human col	345	36	58.1	856	4	ABB59118	Abb59118 Drosophil
273	36	58.1	306	6	ABU03741	Abu03741 Human exp	346	36	58.1	857	8	ADY04305	Ady04305 Plant ful
274	36	58.1	319	4	AAW71967	Aaw71967 Human olf	347	36	58.1	948	8	ADU37227	Adu37227 Human nuc
275	36	58.1	331	5	AAW71348	Aaw71348 Human GPC	348	36	58.1	1001	9	ADY25724	Ady25724 MPTEN RNF
276	36	58.1	331	7	ADC86043	Adc86043 Human GPC	349	36	58.1	1015	8	ADU55041	Adu55041 Human pro
277	36	58.1	332	8	ADU00337	Adu00337 Iron hydr	350	36	58.1	1015	8	ADU18078	Adu18078 Human can
278	36	58.1	334	8	ADX95723	Adx95723 Plant ful	351	36	58.1	1050	4	ABB44551	Abb44551 Human wou
279	36	58.1	343	5	ADH48922	Adh48922 NOV87 pro	352	36	58.1	1066	4	ABG15870	Abg15870 Novel hum
280	36	58.1	343	5	ADC39102	Adc39102 Novel hum	353	36	58.1	1170	8	ADJ76124	Adj76124 Marker ge
281	36	58.1	357	8	ADN22765	Adn22765 Bacterial	354	36	58.1	1173	8	ADR08891	Adr08891 Human pro
282	36	58.1	412	7	ADH88781	Adh88781 Enterococ	355	36	58.1	1278	4	AAW62030	Aaw62030 Recombina
283	36	58.1	412	8	ADX73888	Adx73888 Plant ful	356	36	58.1	1278	5	ABP65735	Abp65735 Bifidobac
284	36	58.1	412	8	ADY13137	Ady13137 Plant ful	357	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
285	36	58.1	413	8	ADK66729	Adk66729 Plant ful	358	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
286	36	58.1	419	4	ABG22112	Abg22112 Novel hum	359	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
287	36	58.1	420	8	ADN19118	Adn19118 Bacterial	360	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
288	36	58.1	430	2	AAW26173	Aaw26173 Part of M	361	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
289	36	58.1	435	2	AAW98458	Aaw98458 H. pylori	362	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
290	36	58.1	435	6	ABU71065	Abu71065 Human adi	363	36	58.1	1523	6	ABP65735	Abp65735 Bifidobac
291	36	58.1	453	4	AAW99944	Aaw99944 ERA bindi	364	35.5	57.3	3321	8	ADQ19071	Adq19071 Human sof
292	36	58.1	454	7	ADC31121	Adc31121 Human nov	365	35.5	57.3	3321	8	ADQ19071	Adq19071 Human sof
293	36	58.1	454	8	ADY07484	Ady07484 Plant ful	366	35.5	57.3	3321	8	ADQ19071	Adq19071 Human sof
294	36	58.1	474	6	ABU42664	Abu42664 Protein e	367	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
295	36	58.1	478	5	ABP40220	Abp40220 Staphyloc	368	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
296	36	58.1	478	6	ADU07285	Adu07285 Staphyloc	369	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
297	36	58.1	485	6	ABU24811	Abu24811 Protein e	370	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
298	36	58.1	491	2	AAW09771	Aaw09771 Pichia ac	371	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
299	36	58.1	496	8	ADY23231	Ady23231 Plant ful	372	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
300	36	58.1	509	2	AAW30429	Aaw30429 Major Yo	373	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
301	36	58.1	524	4	ABBS5149	Abbs5149 Drosophil	374	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
302	36	58.1	547	8	ADY08973	Ady08973 Plant ful	375	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
303	36	58.1	560	2	AAW26763	Aaw26763 Human DNA	376	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
304	36	58.1	560	3	AAW71054	Aaw71054 Human Hec	377	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
305	36	58.1	560	6	ABU03739	Abu03739 Human exp	378	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
306	36	58.1	560	6	ABU03740	Abu03740 Human exp	379	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
307	36	58.1	560	6	ABU03744	Abu03744 Human exp	380	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
308	36	58.1	560	6	ABU03743	Abu03743 Human exp	381	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
309	36	58.1	560	6	ABU03745	Abu03745 Human exp	382	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
310	36	58.1	560	6	ABU03742	Abu03742 Human exp	383	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
311	36	58.1	560	7	ADJ68587	Adj68587 Human hea	384	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
312	36	58.1	560	8	ADH13195	Adh13195 Human mal	385	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
313	36	58.1	560	8	ADL70662	Adl70662 Human Cdc	386	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
314	36	58.1	560	8	ADQ31120	Adq31120 N-end rul	387	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
315	36	58.1	560	8	ADQ31121	Adq31121 N-end rul	388	35	56.5	3321	8	ADQ19071	Adq19071 Human sof
316	36	58.1	560	8	ADO19410	Ado19410 Human PRO	389	35	56.5	3321	8	ADQ19071	Adq19071 Human sof

390	35	56.5	112	6	ADA41371	AdA41371 Human sec	463	35	56.5	305	8	ADK01186	AdK01186 DNA polym
391	35	56.5	112	7	ADD38057	AdD38057 Human sec	464	35	56.5	305	8	ADJ79405	AdJ79405 A. aeolic
392	35	56.5	112	7	ADD38056	AdD38056 Human sec	465	35	56.5	305	8	ADJ84845	AdJ84845 A. aeolic
393	35	56.5	112	8	ADL77241	AdL77241 Albumin f	466	35	56.5	305	8	ADM77633	AdM77633 DNA polym
394	35	56.5	112	8	ADL77245	AdL77245 Albumin f	467	35	56.5	305	8	ADM66300	AdM66300 A. aeolic
395	35	56.5	122	8	ADM87337	AdM87337 Human pro	468	35	56.5	305	8	ADO04353	AdO04353 A. aeolic
396	35	56.5	123	4	AAE04203	AAE04203 Human gen	469	35	56.5	305	8	ADP82430	AdP82430 Aquifex a
397	35	56.5	123	4	AAE04221	AAE04221 Human gen	470	35	56.5	305	8	ADP15637	AdP15637 A. aeolic
398	35	56.5	123	5	ABG64480	ABG64480 Human alb	471	35	56.5	305	9	ADY55128	AdY55128 A. aeolic
399	35	56.5	123	5	ABG664479	ABG664479 Human alb	472	35	56.5	305	9	ADZ76704	AdZ76704 A. aeolic
400	35	56.5	123	8	ADL77746	AdL77746 Albumin f	473	35	56.5	305	9	AEA24712	AEa24712 Aquifex a
401	35	56.5	123	8	ADL77747	AdL77747 Albumin f	474	35	56.5	305	9	AEA34135	AEa34135 Aquifex a
402	35	56.5	123	8	ABM81542	ABm81542 Tumour-as	475	35	56.5	305	9	ABE51539	ABe51539 DNA polym
403	35	56.5	123	8	ADS11034	AdS11034 Human the	476	35	56.5	310	6	ABU44214	ABu44214 Protein e
404	35	56.5	125	3	ABG10708	ABg10708 Arabidops	477	35	56.5	315	4	AAU40846	AAu40846 Human pol
405	35	56.5	139	3	ABG28273	ABg28273 Arabidops	478	35	56.5	318	9	ADM17393	AdM17393 Eucalyptu
406	35	56.5	139	4	AAE01703	AAE01703 Human gen	479	35	56.5	322	7	ADE61776	AdE61776 Rat Prote
407	35	56.5	139	4	AAE03783	AAE03783 Human gen	480	35	56.5	322	7	ADE61772	AdE61772 Rat Prote
408	35	56.5	139	5	ABG63979	ABg63979 Human alb	481	35	56.5	332	6	ABU30483	ABu30483 Protein e
409	35	56.5	139	5	ABG65498	ABg65498 Human alb	482	35	56.5	333	7	ABM89581	ABm89581 Rice abio
410	35	56.5	139	6	ADA57494	AdA57494 Human sec	483	35	56.5	336	3	AAU79181	AAu79181 Haematopo
411	35	56.5	139	6	ADA41372	AdA41372 Human sec	484	35	56.5	336	3	ABU33155	ABu33155 Protein e
412	35	56.5	139	6	ADD38058	AdD38058 Human sec	485	35	56.5	336	9	ABE42108	ABe42108 L. pneumo
413	35	56.5	139	8	ADL77244	AdL77244 Albumin f	486	35	56.5	337	5	ABE93560	ABe93560 Herbicida
414	35	56.5	139	8	ADL78765	AdL78765 Albumin f	487	35	56.5	337	5	ABE38948	ABe38948 L. pneumo
415	35	56.5	151	3	ABG28272	ABg28272 Arabidops	488	35	56.5	338	6	ABP78661	ABp78661 N. gonorr
416	35	56.5	153	6	ABU40806	ABu40806 Protein e	489	35	56.5	343	3	AAU53835	AAu53835 Arabidops
417	35	56.5	157	8	ADO43414	AdO43414 Lutzomyia	490	35	56.5	354	3	AAU53826	AAu53826 Arabidops
418	35	56.5	159	8	ADO43396	AdO43396 Lutzomyia	491	35	56.5	354	3	AAU53834	AAu53834 Arabidops
419	35	56.5	168	7	ADFC96032	ADfC96032 E. faeciu	492	35	56.5	362	6	ABU25256	ABu25256 Protein e
420	35	56.5	177	7	ADP07335	ADp07335 Bacterial	493	35	56.5	363	4	AAE03764	AAE03764 Human gen
421	35	56.5	182	4	ABB69904	ABb69904 Drosophil	494	35	56.5	363	5	ABG65499	ABg65499 Human alb
422	35	56.5	182	8	ADS75086	AdS75086 Fruit fly	495	35	56.5	363	6	ADA57126	AdA57126 Human sec
423	35	56.5	187	3	ABG28271	ABg28271 Arabidops	496	35	56.5	363	6	ADA40987	AdA40987 Human sec
424	35	56.5	195	2	AAU56657	AAu56657 Protein p	497	35	56.5	363	7	ADD37945	AdD37945 Human sec
425	35	56.5	195	3	ABG54259	ABg54259 Lactococc	498	35	56.5	363	8	ADL78766	AdL78766 Albumin f
426	35	56.5	196	3	ABG17355	ABg17355 Arabidops	499	35	56.5	369	6	ABU19161	ABu19161 Pachogen
427	35	56.5	196	3	ABG53840	ABg53840 Arabidops	500	35	56.5	369	8	ADK16419	AdK16419 Nanoarcha
428	35	56.5	197	5	ABE55228	ABe55228 Lactococc							
429	35	56.5	199	8	ADQ66247	ADq66247 Novel hum							
430	35	56.5	199	9	ADY64776	AdY64776 S. mansoni							
431	35	56.5	202	7	ADC86733	AdC86733 Human GPC							
432	35	56.5	206	3	ABG10707	ABg10707 Arabidops							
433	35	56.5	211	8	ADY12150	AdY12150 Plant ful							
434	35	56.5	218	3	ABG10706	ABg10706 Arabidops							
435	35	56.5	218	3	ABG17354	ABg17354 Arabidops							
436	35	56.5	218	3	ABG53839	ABg53839 Arabidops							
437	35	56.5	220	8	ADY09116	AdY09116 Plant ful							
438	35	56.5	230	2	AAU69629	AAu69629 B. burgdo							
439	35	56.5	235	6	ABU29765	ABu29765 Protein e							
440	35	56.5	237	5	ABU51848	ABu51848 Helicobac							
441	35	56.5	239	4	ABU70310	ABu70310 Drosophil							
442	35	56.5	245	3	AAU95011	AAu95011 Human sec							
443	35	56.5	247	4	AAU79624	AAu79624 Corynebac							
444	35	56.5	247	6	ADZ22167	AdZ22167 Bacterial							
445	35	56.5	253	8	ABU71069	ABu71069 Human adi							
446	35	56.5	255	5	ABG70639	ABg70639 Human ser							
447	35	56.5	257	8	ADP04637	AdP04637 Sea squir							
448	35	56.5	270	3	ABG53838	ABg53838 Arabidops							
449	35	56.5	276	9	ADW17307	ADw17307 E. grandis							
450	35	56.5	277	5	AAU93085	AAu93085 Arabidops							
451	35	56.5	277	7	ADC46651	AdC46651 Thalecres							
452	35	56.5	277	7	ADC30967	AdC30967 Plant yie							
453	35	56.5	277	8	ADL18483	AdL18483 Plant tra							
454	35	56.5	277	8	ADO02515	AdO02515 Thalecres							
455	35	56.5	277	9	AEA26433	AEa26433 Stress to							
456	35	56.5	288	9	ADW17382	ADw17382 Eucalyptu							
457	35	56.5	288	6	ABU18362	ABu18362 Protein e							
458	35	56.5	298	6	ADK36187	AdK36187 Acinetoba							
459	35	56.5	299	8	ADK92360	AdK92360 Plant ful							
460	35	56.5	304	5	ABB84806	ABb84806 Aquifex a							
461	35	56.5	305	8	ADJ67896	AdJ67896 A. aeolic							
462	35	56.5	305	8	ADJ68108	AdJ68108 A. aeolic							

ALIGNMENTS

RESULT 1

ADRA47262
ID ADRA47262 standard; peptide, 13 AA.

XX AC ADRA47262;

XX DT 18-NOV-2004 (first entry)

XX DE Human alpha2-antiplasmin cleaving enzyme internal sequence aa 487-499.

XX KW antiinflammatory; cytostatic; vulnery; antiarteriosclerotic;
XX KW antithrombotic; vascular; cerebroprotective; pulmonary;
XX KW fibrin-remodeling disorder; inflammatory condition; arthritis;
XX KW organ fibrosis; scarring; cancer; metastasis; atherothrombotic disease;
XX KW coronary artery thrombosis; stroke; pulmonary embolism; thrombosis;
XX KW human; alpha2-antiplasmin cleaving enzyme.

XX OS Homo sapiens.

XX PN WO2004072240-A2.

XX XX 26-AUG-2004.

XX PD 07-FEB-2004; 2004WO-US003398.

XX PF 07-FEB-2003; 2003US-0445774P.

XX PR 06-FEB-2004; 2004US-00774242.

XX XX (MCKE)/ MCKEE P A.

XX PA (LEEK/) LEE K N.

PA (JACK/) JACKSON K W.
 PA (CHRI/) CHRISTIANSEN V J.
 XX
 XX
 PI Mcke PA, Lee KN, Jackson KW, Christiansen VJ;
 XX
 XX WPI; 2004-625848/60.
 DR
 XX
 XX New alpha-2-antiplasmin cleaving enzyme, useful for treating conditions
 PT involving fibrin, e.g. inflammatory conditions such as all forms of
 PT arthritis, organ fibrosis, undesirable scarring, cancer, or
 PT atherothrombotic disease.
 XX
 XX
 PS Claim 2; SEQ ID NO 6; 40pp; English.
 CC
 CC The invention relates to an alpha2-antiplasmin cleaving enzyme comprising
 CC a protein having a molecular weight of 180 kD in a dimeric form as
 CC determined by SDS-PAGE, where each subunit of the dimeric form has a
 CC molecular weight of 97 kD as determined by SDS-PAGE, and where the enzyme
 CC cleaves precursor alpha2-antiplasmin at the Prol2-Asn13 bond. The enzyme,
 CC inhibitors of the enzyme and methods for detecting such inhibitors are
 CC useful for treating conditions involving fibrin, e.g. inflammatory
 CC conditions such as all forms of arthritis, organ fibrosis, undesirable
 CC scarring, cancer or its metastases; or atherothrombotic disease such as
 CC coronary artery thrombosis, stroke, pulmonary embolism, all other forms
 CC of arterial and venous thromboses. This peptide is an internal sequence
 CC of the alpha2-antiplasmin cleaving enzyme corresponding to amino acids
 CC 487-499.
 XX
 XX Sequence 13 AA;
 SQ
 Query Match 100.0%; Score 62; DB 8; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.0065;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILEENKELENALK 13
 Db 1 ILEENKELENALK 13
 |||||
 |||||
 RESULT 2
 AEB94227
 ID AEB94227 standard; protein; 723 AA.
 XX
 AC AEB94227;
 XX
 DT 06-OCT-2005 (first entry).
 XX
 DE Human soluble FAP alpha dimer SEQ ID NO:70.
 XX
 KW immune inhibition; fibroblast activation protein alpha dimer;
 KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
 KW immune disorder; neurological disease; autoimmune disease;
 KW immunosuppressive; graft versus host disease; transplant rejection;
 KW endotoxic shock; osteoarthritis; antiarthritic; osteopathic;
 KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
 KW inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
 KW cardiovascular disease; metabolic disorder; hashimoto's disease;
 KW antithyroid; endocrine disease; inflammatory bowel disease;
 KW antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
 KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
 KW autoimmune hepatitis; antiinflammatory; hepatotropic;
 KW systemic lupus erythematosus; dermatological; dermatological disease;
 KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;
 KW hematological disease; rheumatic fever; antipyretic; Crohns disease;
 KW psoriasis; antipsoriatic; Graves disease; antithyroid;
 KW respiratory syncytial virus infection; respiratory-gen.; virucide;
 KW CD26 dipeptidyl peptidase IV; DPPIV.
 XX
 OS Homo. sapiens.
 OS
 XX
 PN WC2005071073-A1.
 XX
 XX 04-AUG-2005.
 PD

XX 10-JAN-2005; 2005WO-US000709.
 XX
 XX 09-JAN-2004; 2004US-0535577P.
 XX
 XX (POIN-) POINT THERAPEUTICS INC.
 XX
 XX Mclean PA, Jones B, Miller GT, Jesson MI;
 XX
 XX WPI; 2005-564220/57.
 XX
 XX Down-regulating an immune response comprises administering to a subject
 PT in need a fibroblast activation protein (FAP) alpha dimer enzyme in an
 PT amount effective to down-regulate an immune response.
 XX
 XX Claim 134; SEQ ID NO 70; 177pp; English.
 PS
 XX
 XX The invention relates to a method of down-regulating an immune response,
 CC which comprises administering to a subject a fibroblast activation
 CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
 CC an immune response. Also included are the following: a composition
 CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
 CC carrier, where the composition is sterile and lacks an adjuvant; a
 CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
 CC acceptable carrier, and a non-adjuvant second agent; a composition
 CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
 CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
 CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
 CC The method further comprises administering to the subject a second agent.
 CC The second agent is an anti-inflammatory agent, immunosuppressant, or
 CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
 CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
 CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
 CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
 CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
 CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
 CC substitution relative to wild type FAP alpha dimer. The amino acid
 CC substitution is present in the beta-propeller domain, the catalytic
 CC domain, or an N-linked glycosylation site and alters disulfide bond
 CC formation. The immune response is an especially an IL-1 mediated
 CC condition, abnormal immune response selected from inflammation,
 CC autoimmune disease, sepsis, graft versus host disease, transplant
 CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
 CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
 CC response is subsequent to an infection, such as an RSV infection. The
 CC autoimmune disease is selected from c. autoimmune thyroiditis, systemic
 CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
 CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
 CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
 CC sclerosis. The subject does not have cancer or a predisposition to
 CC cancer. The present sequence represents the amino acid sequence of human
 CC soluble fibroblast activation protein (FAP) alpha dimer (corresponding to
 CC AEB94159 minus N-terminal amino acids 1-37).
 XX
 SQ Sequence 723 AA;
 Query Match 100.0%; Score 62; DB 9; Length 723;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILEENKELENALK 13
 Db 450 ILEENKELENALK 462
 |||||
 |||||
 RESULT 3
 AEB94218
 ID AEB94218 standard; protein; 734 AA.
 XX
 AC AEB94218;
 XX
 DT 06-OCT-2005 (first entry)
 XX

CC The invention relates to a method of down-regulating an immune response,
 CC which comprises administering to a subject a fibroblast activation
 CC protein (FAP) alpha dimer enzyme in an amount effective to down-regulate
 CC an immune response. Also included are the following: a composition
 CC comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable
 CC carrier, where the composition is sterile and lacks an adjuvant; a
 CC composition comprising a FAP alpha dimer enzyme in a pharmaceutically
 CC acceptable carrier, and a non-adjuvant second agent; a composition
 CC comprising a FAP alpha dimer enzyme comprising an amino acid substitution
 CC of A657D; and a composition comprising a FAP alpha dimer enzyme lacking
 CC amino acids 269-448 and comprising amino acids 269-448 from mouse FAP.
 CC The method further comprises administering to the subject a second agent.
 CC The second agent is an anti-inflammatory agent, immunosuppressant, or
 CC anti-infective agent such as antibacterial, antiviral, antifungal, anti-
 CC parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is wild
 CC type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation
 CC mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The
 CC FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a
 CC DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid
 CC substitution relative to wild type FAP alpha dimer. The amino acid
 CC substitution is present in the beta-propeller domain, the catalytic
 CC domain, or an N-linked glycosylation site and alters disulfide bond
 CC formation. The immune response is an especially an IL-1 mediated
 CC condition, abnormal immune response selected from inflammation,
 CC autoimmune disease, sepsis, graft versus host disease, transplant
 CC rejection, toxic shock syndrome, allergy, asthma, atherosclerosis,
 CC osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune
 CC response is subsequent to an infection, such as an RSV infection. The
 CC autoimmune disease is selected from C, autoimmune thyroiditis, systemic
 CC lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever,
 CC Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease,
 CC myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple
 CC sclerosis. The subject does not have cancer or a predisposition to
 CC cancer. The present sequence represents the amino acid sequence of human
 CC soluble fibroblast activation protein (FAP) alpha dimer. Note: SEQ ID
 CC NO:3 ABB94160 only encodes amino acids 84 to 750 the present sequence.

XX Sequence 750 AA;

Query Match 100.0%; Score 62; DB 9; Length 750;
 Best Local Similarity 100.0%; Pred. NO. 0.48;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 DB 476 ILEENKELENALK 488

RESULT 5
 ID AAW31963 standard; protein; 759 AA.

XX AAW31963;

XX 21-JAN-1998 (first entry)

XX Human fibroblast activation protein-alpha (FAP-alpha).

XX Fibroblast activation protein-alpha; FAP-alpha; human;
 XX type II integral membrane proteins; epithelial cancer.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1..759
 FT /label= FAP-alpha
 FT /note= "Fibroblast activation protein alpha"

FT Modified-site 49..51
 FT /note= "Asn is N-glycosylated"

FT Region 191..209

FT /label= fap-1

FT /note= "Putative functional element; not defined further
 in specification"

FT

FT Modified-site 214..216
 FT /note= "Asn is N-glycosylated"
 FT Region 220..240
 FT /label= fap-2
 FT /note= "Putative functional element; not defined further
 in specification"
 FT Region 510..521
 FT /label= fap-3
 FT /note= "Putative functional element; not defined further
 in specification"
 FT Region 625..668
 FT /note= "alternative sequence found in AAW27438"
 FT Modified-site 637..639
 FT /note= "Asn is N-glycosylated"
 FT Modified-site 678..680
 FT /note= "Asn is N-glycosylated"
 XX WO9734927-A1.
 XX 25-SEP-1997.
 XX 12-MAR-1997; 97WO-US004215.
 XX 18-MAR-1996; 96US-00619280.
 XX (LUDWIG) LUDWIG INST CANCER RES.
 XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 XX Zimmerman R, Park JB, Rettig W, Old LJ;
 WPI; 1997-480164/44.
 DR N-PSDB; AAT89286.

XX Dimeric fibroblast activation protein alpha - capable of degrading
 extracellular matrix protein, useful to treat epithelial cancers.

XX Example 5; Fig 1; 40pp; English.

XX This sequence represents the dimeric fibroblast activation protein alpha
 (FAP-alpha). The FAP-alpha molecule is capable of degrading extracellular
 matrix proteins and exhibits structural features typical of type II
 integral membrane proteins, including a large COOH-terminal extracellular
 domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.
 CC The putative extracellular domain is described as containing five
 CC potential N-glycosylation sites, eleven conserved cysteine residues and
 CC three segments corresponding to highly conserved catalytic domains
 CC characteristic of serine proteases. Most common types of epithelial
 CC cancers including breast, lung, skin, pancreas and colorectal carcinomas
 CC contain abundant FAP-alpha reactive stromal fibroblasts which accompany
 CC tumour blood vessels. The induction of FAP-alpha fibroblasts at times and
 CC sites of tissue remodelling during foetal development, tissue repair and
 CC carcinogenesis is consistent with its putative role in normal fibroblast
 CC physiology. FAP alpha inhibitors, particularly a collagen derivative or
 CC (S)-valyl-pyrrolidine-2(R)-boronic acid, may be used to treat various
 CC epithelial cancers, e.g. breast, lung, skin, pancreas or colon cancer.
 CC Note: This sequence differs from the amino acid sequence of FAP-alpha
 CC found in AAW27438 between residues 625 to 670

XX Sequence 759 AA;

Query Match 100.0%; Score 62; DB 2; Length 759;
 Best Local Similarity 100.0%; Pred. NO. 0.49;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 487 ILEENKELENALK 499

RESULT 6

AAW27438

ID AAW27438 standard; protein; 760 AA.

XX

```

AC AAW27438;
XX
XX
DT 21-JAN-1998 (first entry)
XX
XX
DE Human fibroblast activation protein-alpha (FAP-alpha).
XX
XX
KW Fibroblast activation protein-alpha; FAP-alpha; human;
XX type II integral membrane proteins; epithelial cancer.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 1..760
FT /label= FAP-alpha
FT /note= "Fibroblast activation protein alpha"
FT 625..670
FT /note= "alternative sequence found in AAW31963"
XX
XX
PN WO9734927-A1.
XX
XX
PD 25-SEP-1997.
XX
XX
PF 12-MAR-1997; 97WO-US004215.
XX
XX
PR 18-MAR-1996; 96US-00619280.
XX
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (BOEH-) BOEHRINGER INGELHEIM INT GMBH.
XX
XX
PI Zimmerman R, Park JE, Rettig W, Old LJ;
XX
XX
DR WPI; 1997-480164/44.
DR N-PSDB; AAT90148.
XX
XX
PT Dimeric fibroblast activation protein alpha - capable of degrading
PT extracellular matrix protein, useful to treat epithelial cancers.
XX
XX
PS Claim 2; Page 25-28; 40pp; English.
XX
XX
CC This sequence represents a dimeric fibroblast activation protein alpha
CC (FAP-alpha) which is capable of degrading extracellular matrix proteins.
CC The FAP-alpha molecule exhibits structural features typical of type II
CC integral membrane proteins, including a large COOH-terminal extracellular
CC domain, a hydrophobic transmembrane segment and a short cytoplasmic tail.
CC Most common types of epithelial cancers including breast, lung, skin,
CC pancreas and colorectal carcinomas contain abundant FAP-alpha reactive
CC stromal fibroblasts which accompany tumour blood vessels. The induction
CC of FAP-alpha fibroblasts at times and sites of tissue remodelling during
CC fetal development, tissue repair and carcinogenesis is consistent with
CC its putative role in normal fibroblast physiology. FAP alpha inhibitors,
CC particularly a collagen derivative or (S)-valyl-pyrrolidine-2(R)-boronic
CC acid, may be used to treat various epithelial cancers, e.g. breast, lung,
CC skin, pancreas or colon cancer. Note: This sequence differs from the
CC amino acid sequence of FAP-alpha found in AAW31963 between residues 625
CC to 668
XX
XX
SQ Sequence 760 AA;
Query Match 100.0%; Score 62; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILEENKELENALK 13
| | | | | | | | | | | | | | |
DB 487 ILEENKELENALK 499

RESULT 7
ABR47452
ID ABR47452 standard; protein; 760 AA.
XX
XX
AC ABR47452;
XX
XX

```

```

DT 12-JUN-2003 (first entry)
XX
XX
DE Breast cancer associated protein sequence SEQ ID NO:136.
XX
XX
KW Human; breast cancer; cytostatic; gene therapy.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2003004989-A2.
XX
XX
PD 16-JAN-2003.
XX
XX
PF 21-JUN-2002; 2002WO-US019669.
XX
XX
PR 21-JUN-2001; 2001US-0299887P.
PR 27-JUN-2001; 2001US-0301572P.
PR 18-JUL-2001; 2001US-0306501P.
PR 25-SEP-2001; 2001US-0325002P.
PR 05-MAR-2002; 2002US-0362585P.
PR 14-MAY-2002; 2002US-0380391P.
XX
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
XX
PI Lillie J, Gannavarapu M, Giatt K, Hoersh S, Kamatkar S;
PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI Bast RC, Hortobagyi GN, Puzstai L, Meric F, Sahin A, Mills GB;
XX
XX
DR WPI; 2003-210381/20.
DR N-PSDB; ACC50144.
XX
XX
PT Breast cancer diagnosis or treatment by comparing the level of expression
PT of a marker in a patient sample with that in the control non-breast
PT cancer sample.
XX
XX
PS Claim 1; SEQ ID NO 136; 128pp; English.
XX
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 760 AA;
Query Match 100.0%; Score 62; DB 6; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILEENKELENALK 13
| | | | | | | | | | | | | | |
DB 487 ILEENKELENALK 499

RESULT 8
ADN95552
ID ADN95552 standard; protein; 760 AA.
XX
XX
AC ADN95552;
XX
XX
DT 01-JUL-2004 (first entry)
XX
XX
DE Human BEC/LEC-related protein sequence SeqID475.
XX
XX
KW growth; differentiation; blood endothelial cell; BEC;
XX lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
XX

```


KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX Homo sapiens.
XX WO2003080640-A1.
XX 02-OCT-2003.
XX 07-MAR-2003; 2003WO-US006900.
XX 07-MAR-2002; 2002US-0363019P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX (LICN) LICENTIA LTD.
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX WPI; 2003-876899/81.
XX N-PSDB; ADN95553.
XX Example 1; SEQ ID NO 475; 176pp; English.
XX This invention relates to a method of differentially modulating the
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC endothelial cells (LEC) comprising contacting endothelial cells with a
CC composition comprising an agent that differentially modulates blood or
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC identifying a human subject with lymphoedema and with a mutation in at
CC least one allele of a gene encoding a LEC protein, where the mutation
CC correlates with lymphoedema in human subjects, and with the proviso that
CC the LEC protein is not VEGFR-3; and administering to the subject a
CC composition comprising a lymphatic growth agent selected from VEGF-C or
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC the development of compounds with an antiangiogenic, cytostatic,
CC vasotrophic or antiinflammatory activity or for gene therapy. The method
CC is useful in modulating the growth or differentiation of blood
CC endothelial cells or lymphatic endothelial cells, in treating hereditary
CC lymphoedema, in screening for an endothelial cell disorder or
CC predisposition to the disorder or in monitoring the efficacy or toxicity
CC of a drug on endothelial cells. The agent is useful in manufacturing a
CC medicament for the differential modulation of blood vessel endothelial
CC cell or lymphatic vessel endothelial cell growth or differentiation. The
CC lymphatic growth agent may also be used in manufacturing a medicament for
CC the treatment of hereditary lymphoedema resulting from a mutation in a
CC LEC gene or of other diseases involving the lymphatic vessels, such as
CC various inflammatory diseases and cancer metastasis via the lymphatic
CC system. The present sequence is that of a human LEC/BEC differentially
CC expressed protein which is related to the method of the invention. Note:
CC This sequence does not appear in the specification but was obtained by
CC the indexer using the source data given in table 14 of the specification.
XX SQ Sequence 760 AA;
Query Match 100.0%; Score 62; DB 7; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 487 ILEENKELENALK 499
RESULT 9
ID ADQ21351 standard; protein; 760 AA.
XX ADQ21351;
XX ADQ21351;
XX 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated protein - SEQ ID 4171.
DE

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX Homo sapiens.
XX WO2004048938-A2.
XX 10-JUN-2004.
XX 26-NOV-2003; 2003WO-US038193.
XX 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX Example 2; SEQ ID NO 4171; 210pp; English.
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX protein of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 760 AA;
Query Match 100.0%; Score 62; DB 8; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.49; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 487 ILEENKELENALK 499
RESULT 10
ADW14775
ID ADW14775 standard; protein; 760 AA.
XX ADW14775;
XX 07-APR-2005 (first entry)
XX Tumor-associated antigenic target TAT430.
XX Tumor-associated antigenic target; TAT430; cancer; neoplasm; cytostatic;
XX ovary tumor; uterus tumor.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..28
XX Modified-site /label= Signal_peptide
XX /note= "N-myristoylation site"
XX Protein 29..760
XX /label= Mature_protein
FT

FT	Region	42..548	/note= "Dipeptidyl peptidase IV"
FT	Modified-site	49..52	/note= "N-glycosylation site"
FT	Modified-site	84..89	/note= "N-myristoylation site"
FT	Modified-site	92..95	/note= "N-glycosylation site"
FT	Modified-site	99..102	/note= "N-glycosylation site"
FT	Modified-site	159..164	/note= "N-myristoylation site"
FT	Modified-site	285..290	/note= "N-myristoylation site"
FT	Modified-site	314..317	/note= "N-glycosylation site"
FT	Modified-site	366..374	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	402..410	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	419..422	/note= "Amidation site"
FT	Region	438..443	/note= "Cytochrome c family heme-binding site signature"
FT	Modified-site	470..475	/note= "N-myristoylation site"
FT	Modified-site	542..547	/note= "N-myristoylation site"
FT	Region	552..629	/note= "Prolyl oligopeptidase family homology"
FT	Modified-site	566..571	/note= "N-myristoylation site"
FT	Active-site	599..629	/note= "Prolyl endopeptidase family serine active site"
FT	Modified-site	626..631	/note= "N-myristoylation site"
FT	Modified-site	636..639	/note= "Glycosaminoglycan attachment site"
FT	Modified-site	639..644	/note= "N-myristoylation site"
FT	Modified-site	679..682	/note= "N-glycosylation site"
FT	Modified-site	738..743	/note= "N-myristoylation site"
XX		W02005003154-A2.	
PN		13-JAN-2005.	
FD		02-JUL-2004; 2004WO-US021353.	
XX		02-JUL-2003; 2003US-0484959P.	
PR		(GETH) GENENTECH INC.	
XX		Ashkenazi A, Cairns B, Dowd P, Frantz G, French D, Gonzalez L;	
PI		Polakis P, Smith V, Wolf B, Wu TD, Zhang Z;	
XX		WPI; 2005-091764/10.	
DR		N-PSDB; ADW14768.	
XX		New Tumor-associated Antigenic Target polypeptides useful for diagnosing	
PT		preventing or treating cell proliferative disorders such as cancer or	
PT		tumor.	
XX		Claim 1; SEQ ID NO 13; 141pp; English.	
PS		The present sequence is that of novel tumor-associated antigenic target	
XX		polypeptide TAT430. TAT polypeptides have been identified ADW14770-	
CC		ADW14776 whose expression is upregulated in particular tumor tissues.	
CC		Tissue expression profiling using Genexpress indicated that TAT430 is	
CC		upregulated in breast, pancreatic, colon, adrenal, bladder, kidney,	
CC		liver, lung, lymphoid, ovarian, prostate and stomach tumor stroma as	

PS Disclosure; SEQ ID NO 2; 177pp; English.

XX The invention relates to a method of down-regulating an immune response, which comprises administering to a subject a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response. Also included are the following: a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, where the composition is sterile and lacks an adjuvant; a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, and a non-adjuvant second agent; a composition comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising a FAP alpha dimer enzyme lacking amino acids 269-448 and comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, anti-parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is a wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of human wild type fibroblast activation protein (FAP) alpha dimer.

XX Sequence 760 AA;

Query Match 100.0%; Score 62; DB 9; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||||

Db 487 ILEENKELENALK 499

RESULT 12

ID AEB94163 standard; protein; 761 AA.

AC AEB94163;

XX AEB94163;

DT 06-OCT-2005 (first entry)

XX Mouse wild type fibroblast activation protein (FAP) alpha dimer.

DE Immune inhibition; fibroblast activation protein alpha dimer;
KW FAP alpha dimer; guillain barre syndrome; antiinflammatory; cns-gen.;
KW immune disorder; neurological disease; autoimmune disease;
KW immunosuppressive; graft versus host disease; transplant rejection;
KW endotoxin shock; osteoarthritis; antiarthritic; osteopathic;
KW musculoskeletal disease; allergy; antiallergic; asthma; antiasthmatic;
KW inflammation; respiratory disease; atherosclerosis; antiarteriosclerotic;
KW cardiovascular disease; metabolic disorder; hashimoto's disease;
KW antithyroid; endocrine disease; inflammatory bowel disease;
KW antiinflammatory; gastrointestinal-gen.; gastrointestinal disease;
KW rheumatoid arthritis; antirheumatic; multiple sclerosis; neuroprotective;
KW autoimmune hepatitis; antiinflammatory; hepatotropic;
KW systemic lupus erythematosus; dermatological; dermatological disease;
KW uveitis; ophthalmological; autoimmune hemolytic anemia; antianemic;

KW hematological disease; rheumatic fever; antipyretic; Crohn's disease;
KW psoriasis; antiparasitic; graves disease; antithyroid;
KW respiratory syncytial virus infection; respiratory-gen.; virucide;
enzyme.
XX Mus musculus.
XX WO2005071073-A1.
XX 04-AUG-2005.
XX 10-JAN-2005; 2005WO-US000709.
XX 09-JAN-2004; 2004US-0535577P.
XX (POIN-) POINT THERAPEUTICS INC.
XX Mclean PA, Jones B, Miller GT, Jenson MI;
PI WPI; 2005-564220/57.
XX N-PSDB; AEB94162.
PT Down-regulating an immune response comprises administering to a subject in need a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response.
PT Disclosure; SEQ ID NO 6; 177pp; English.
XX The invention relates to a method of down-regulating an immune response, which comprises administering to a subject a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response. Also included are the following: a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, where the composition is sterile and lacks an adjuvant; a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, and a non-adjuvant second agent; a composition comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, anti-parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is a wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of human wild type fibroblast activation protein (FAP) alpha dimer.

QY 1 ILEENKELENALK 13
:|||||

Db 487 VLEENKELENSLR 499

Query Match 88.7%; Score 55; DB 9; Length 761;
Best Local Similarity 76.9%; Pred. No. 6;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

KW hematological disease; rheumatic fever; antipyretic; Crohn's disease;
KW psoriasis; antiparasitic; graves disease; antithyroid;
KW respiratory syncytial virus infection; respiratory-gen.; virucide;
enzyme.
XX Mus musculus.
XX WO2005071073-A1.
XX 04-AUG-2005.
XX 10-JAN-2005; 2005WO-US000709.
XX 09-JAN-2004; 2004US-0535577P.
XX (POIN-) POINT THERAPEUTICS INC.
XX Mclean PA, Jones B, Miller GT, Jenson MI;
PI WPI; 2005-564220/57.
XX N-PSDB; AEB94162.
PT Down-regulating an immune response comprises administering to a subject in need a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response.
PT Disclosure; SEQ ID NO 6; 177pp; English.
XX The invention relates to a method of down-regulating an immune response, which comprises administering to a subject a fibroblast activation protein (FAP) alpha dimer enzyme in an amount effective to down-regulate an immune response. Also included are the following: a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, where the composition is sterile and lacks an adjuvant; a composition comprising a FAP alpha dimer enzyme in a pharmaceutically acceptable carrier, and a non-adjuvant second agent; a composition comprising a FAP alpha dimer enzyme comprising an amino acid substitution of A657D; and a composition comprising amino acids 269-448 from mouse FAP. The method further comprises administering to the subject a second agent. The second agent is an anti-inflammatory agent, immunosuppressant, or anti-infective agent such as antibacterial, antiviral, antifungal, anti-parasitic or anti-mycobacterial agent. The FAP alpha dimer enzyme is a wild type FAP alpha dimer enzyme. The FAP alpha dimer enzyme is a truncation mutant. The FAP alpha dimer enzyme is a fusion or chimeric protein. The FAP alpha dimer enzyme is a heterodimer of a FAP alpha monomer and a DPPIV/CD26 monomer. The FAP alpha dimer enzyme comprises an amino acid substitution relative to wild type FAP alpha dimer. The amino acid substitution is present in the beta-propeller domain, the catalytic domain, or an N-linked glycosylation site and alters disulfide bond formation. The immune response is an especially an IL-1 mediated condition, abnormal immune response selected from inflammation, autoimmune disease, sepsis, graft versus host disease, transplant rejection, toxic shock syndrome, allergy, asthma, atherosclerosis, osteoarthritis, and Guillain-Barre's syndrome. The abnormal immune response is subsequent to an infection, such as an RSV infection. The autoimmune disease is selected from c, autoimmune thyroiditis, systemic lupus erythematosus (SLE), uveitis, hemolytic anemias, rheumatic fever, Crohn's disease, Guillain-Barre's syndrome, psoriasis, Graves' disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis and multiple sclerosis. The subject does not have cancer or a predisposition to cancer. The present sequence represents the amino acid sequence of mouse wild type fibroblast activation protein (FAP) alpha dimer.

RESULT 13
ADX90113
ID ADX90113 standard; protein; 562 AA.
XX
AC ADX90113;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 52777.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAK/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 52777; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 562 AA;
Query Match 79.0%; Score 49; DB 8; Length 562;
Best Local Similarity 69.2%; Pred. No. 37;

Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ILEENKELENALK 13
:||||: |||||;
Db 1 VLEENQSLNALR 13
RESULT 14
AAW20957
ID AAW20957 standard; protein; 223 AA.
XX
AC AAW20957;
XX
DT 21-JUL-1997 (first entry)
XX
DE H. pylori cytoplasmic protein hpi13922orf30.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
XX
OS Helicobacter pylori.
XX
PN WO9640893-A1.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009122.
XX
PR 07-JUN-1995; 95US-00487032.
PR 01-APR-1996; 96US-00630405.
XX
PA (ASTR) ASTRA AB.
XX
PI Smith D, Berglindh OT, Mellgaard BL;
XX WPI; 1997-052306/05.
XX
DR N-PSDB; AAT68210.
XX
PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) -
PT useful for vaccines to treat or prevent H. pylori infection, and to
PT detect Helicobacter.
XX
PS Claim 61; Page 1349; 1481pp; English.
XX
CC The present sequence represents a Helicobacter pylori cytoplasmic protein
CC that may be used in a vaccine to prevent or treat H. pylori infection or
CC to identify H. pylori polypeptide binding compounds, useful as potential
CC H. pylori life cycle activators or inhibitors. The genomic sequence of H.
CC pylori (ATCC 55679) was determined from overlapping contigs generated by
CC mechanically shearing the bacterial DNA. The sequences were analysed for
CC ORF of at least 180 nucleotides, and the predicted coding regions defined
CC by computer evaluation. To identify likely H. pylori antigens for vaccine
CC development, the amino acid sequences predicted from various ORF were
CC analysed for significant homology to other known or exported membrane
CC proteins. Having identified and determined the sequences of interest,
CC particular regions can be isolated from H. pylori by PCR amplification
CC for recombinant polypeptide production, e.g. in E. coli hosts
XX
SQ Sequence 223 AA;
Query Match 69.4%; Score 43; DB 2; Length 223;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ILEENKELENALK 13
|||:|||||;
Db 208 ILTESKIEELALK 220
RESULT 15
AAW20427
ID AAW20427 standard; protein; 296 AA.

```

XX AC AAW20427;
XX DT 14-JUL-1997 (first entry)
XX DE H. pylori cytoplasmic protein 3409087.aa.
XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
XX KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
XX KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; ds.
XX OS Helicobacter pylori.
XX FH Key Location/Qualifiers
XX FT Misc-difference 285
XX FT /note= "encoded by TGS"
XX PN WO9640893-A1.
XX PD 19-DEC-1996.
XX PF 06-JUN-1996; 96WO-US009122.
XX PR 07-JUN-1995; 95US-00487032.
XX PR 01-APR-1996; 96US-00630405.
XX PA (ASTR ) ASTRA AB.
XX PI Smith D, Berglindh OT, Mellgaard BL;
XX DR WPI; 1997-052306/05.
XX DR N-PSDB; AAT67602.
XX PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) -
XX PT useful for vaccines to treat or prevent H. pylori infection, and to
XX PT detect Helicobacter.
XX PS Claim 61; Page 605; 1481pp; English.
XX CC The present sequence encodes a Helicobacter pylori cytoplasmic protein
XX CC that may be used in a vaccine to prevent or treat H. pylori infection or
XX CC to identify H. pylori polypeptide binding compounds, useful as potential
XX CC H. pylori life cycle activators or inhibitors. The genomic sequence of H.
XX CC pylori (ATCC 55679) was determined from overlapping contigs generated by
XX CC mechanically shearing the bacterial DNA. The sequences were analysed for
XX CC ORF of at least 180 nucleotides, and the predicted coding regions defined
XX CC by computer evaluation. To identify likely H. pylori antigens for vaccine
XX CC development, the amino acid sequences predicted from various ORF were
XX CC analysed for significant homology to other known or exported membrane
XX CC proteins. Having identified and determined the sequences of interest,
XX CC particular regions can be isolated from H. pylori by PCR amplification
XX CC for recombinant polypeptide production, e.g. in E. coli hosts
XX SQ Sequence 296 AA;
XX Query Match 69.4%; Score 43; DB 2; Length 296;
XX Best Local Similarity 69.2%; Pred. No. 1.6e+02;
XX Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 ILEENKELENALK 13
XX DB 176 ILTESKEIEEALK 188
XX RESULT 16
XX ABUS1130
XX ID ABUS1130 standard; protein; 345 AA.
XX AC ABUS1130;
XX XX
XX DT 07-MAY-2003 (first entry)
XX XX Helicobacter pylori selected interacting domain (SID) protein #473.
XX DE

XX KW Protein-protein interaction; ulcer; selected interacting domain; SID.
XX OS Helicobacter pylori.
XX PN WO200266501-A2.
XX PD 29-AUG-2002.
XX PF 28-DEC-2001; 2001WO-EP015428.
XX PR 02-JAN-2001; 2001US-0259302P.
XX XX (HYBR-) HYBRIGENICS.
XX PA (INSP ) INST PASTEUR.
XX PI Legrain P, Rain J, Colland P, De Reuse H, Labigne A;
XX XX WPI; 2002-674910/72.
XX DR N-PSDB; ABX65874.
XX PT New complexes of protein-protein interactions in Helicobacter pylori,
XX PT useful for identifying modulating compounds for treating or preventing
XX PT ulcers in mammals.
XX PS Claim 6; Page 200; 642pp; English.
XX CC The invention describes a complex of protein-protein interactions in
XX CC Helicobacter pylori selected from 421 complexes given in the
XX CC specification. The complex of protein-protein interactions are useful for
XX CC screening for agents which modulate the interaction of proteins.
XX CC Modulating compounds which binds to a targeted bacterial protein may be
XX CC used for treating or preventing ulcers in a human or animal. This is the
XX CC amino acid sequence of a selected interacting domain (SID), identified
XX CC via protein-protein interactions. Note: Where the patent number printed
XX CC at the top of the pages in the specification has obscured areas of
XX CC protein sequence, the indexer has replaced the residue with an X to
XX CC represent an illegible residue
XX SQ Sequence 345 AA;
XX Query Match 69.4%; Score 43; DB 5; Length 345;
XX Best Local Similarity 69.2%; Pred. No. 1.9e+02;
XX Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 ILEENKELENALK 13
XX DB 157 ILTESKEIEEALK 169
XX RESULT 17
XX AAB52523
XX ID AAB52523 standard; protein; 346 AA.
XX XX
XX AC AAB52523;
XX DT 23-FEB-2001 (first entry)
XX DE Helicobacter pylori bait polypeptide #41.
XX KW Helicobacter pylori; two-hybrid system; protein-protein interaction;
XX KW bait polypeptide; gastric ulcer; antibacterial.
XX OS Helicobacter pylori.
XX XX
XX PN WO200066722-A1.
XX PD 09-NOV-2000.
XX XX
XX PF 14-APR-2000; 2000WO-IB000603.
XX PR 30-APR-1999; 99EP-00401066.
XX XX

```

(HYBR-) HYBRIGENICS SA.
 PI Legrain P, Selig L, Rain J;
 XX
 DR WPI; 2000-687535/67.
 DR N-PSDB; AAC97269.
 XX
 PT A two-hybrid system for identifying compounds useful in the treatment of
 PT e.g. gastric ulcers comprises producing a collection of recombinant cell
 PT clones.
 XX
 PS Example 5; Page 131-132; 267pp; English.
 XX
 CC The present sequence is a bait polypeptide used in a Helicobacter pylori
 CC two-hybrid screen to identify protein-protein interactions. The method is
 CC used to identify a recombinant cell clone expressing a prey polypeptide
 CC which is capable of interacting with the bait polypeptide. The two hybrid
 CC system is useful for screening compounds for antibacterial activity. It
 CC may be used in the treatment of gastric ulcers. The polynucleotides are
 CC useful as amplification primers or specific detection probes. The
 CC polypeptides, vectors or host cells can be used as immunogens to produce
 CC mono- or polyclonal antibodies. The polynucleotides, polypeptides,
 CC antibodies, vectors, host cells or modulating agents can be used to
 CC produce a pharmaceutical composition
 XX
 SQ Sequence 346 AA;
 Query Match 69.4%; Score 43; DB 3; Length 346;
 Best Local Similarity 69.2%; Pred. No. 1.9e+02;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ILEENKELENALK 13
 DB 157 ILTESKETSEALK 169
 ||| :||: |||
 ||| :||: |||
 RESULT 18
 ABR53675
 ID ABR53675 standard; protein; 706 AA.
 AC ABR53675;
 XX
 DT 20-JUN-2003 (first entry)
 XX
 DE Protein sequence #SEQ ID 2215.
 XX
 KW Multiprotein complex; eukaryote; drug target; diagnosis.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN EP1258494-A1.
 XX
 PD 20-NOV-2002.
 XX
 PF 20-DEC-2001; 2001EP-00130253.
 XX
 PR 15-MAY-2001; 2001EP-00111774.
 XX
 PS (CELL-) CELLZONE AG.
 XX
 PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
 PI Marzloch M, Schultz JD, Superti-Furga GD;
 XX
 DR WPI; 2003-250078/25.
 DR N-PSDB; ACC61717.
 XX
 PT New isolated protein complexes useful for diagnosing a disease or
 PT disorder, or as a target for an active agent of a pharmaceutical,
 PT preferably a drug target in the treatment or prevention of disease or
 PT disorder.
 XX
 PS Disclosure; SEQ ID NO 2215; 17pp + Sequence Listing; English.
 XX

The invention relates to multiprotein complexes from eukaryotes. Proteins
 of the invention and DNA sequences encoding them are given in records
 ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
 obtainable by using a protein as a bait and isolating the set of proteins
 which is attached thereto from cells. Such protein complexes may comprise
 up to 30 distinct proteins. Protein complexes of the invention are useful
 for diagnosing a disease or disorder, or as a target for an active agent
 of a pharmaceutical, preferably a drug target in the treatment or
 prevention of a disease or disorder. Note: The sequence data for this
 patent is not represented in the printed specification, but is based on
 sequence information supplied by the European Patent Office. The complete
 document is available on CD-ROM
 XX
 SQ Sequence 706 AA;
 Query Match 69.4%; Score 43; DB 6; Length 706;
 Best Local Similarity 81.8%; Pred. No. 4e+02;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LEENKELENAL 12
 DB 22 LEENKELENAL 32
 ||||| |||
 ||||| |||
 RESULT 19
 ADB96560
 ID ADB96560 standard; protein; 706 AA.
 XX
 AC ADB96560;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Yeast histone deacetylase SC_HDA1 protein.
 XX
 KW cytosolic; histone deacetylase; diagnosis; breast; bladder; colon; lung;
 KW rectal; endometrial; ovarian; prostate; cancer; Paget's disease;
 KW melanoma; adenocarcinoma; enzyme.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN WO2002102323-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 14-JUN-2002; 2002WO-US019560.
 XX
 PR 14-JUN-2001; 2001US-0298296P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Jackson DG, Lorenzi MV, Attar RM, Gottardis M;
 XX
 DR WPI; 2003-167442/16.
 XX
 PT New isolated human histone deacetylase polypeptide, useful for diagnosing
 PT and treating disorders with aberrant expression or activity of the
 PT histone deacetylases, such as breast, colon, lung and prostate cancer.
 XX
 PS Disclosure; Fig 14B; 206pp; English.
 XX
 CC The invention relates to the isolation of histone deacetylase
 CC polypeptides and the genes encoding them. The compositions of the present
 CC invention are useful for the diagnosis and treatment of disorders
 CC associated with the aberrant expression or activity of the histone
 CC deacetylases, such as breast, bladder, colon, lung, rectal, endometrial,
 CC ovarian and prostate cancer, Paget's disease, melanoma and
 CC adenocarcinomas. This sequence represents the histone deacetylase SC_HDA1
 CC protein used for sequence comparison with the novel proteins.
 XX
 SQ Sequence 706 AA;
 Query Match 69.4%; Score 43; DB 7; Length 706;
 Best Local Similarity 81.8%; Pred. No. 4e+02;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LEENKELENAL 12
Db 22 LEENKEENSL 32
||||| |.:

RESULT 20

ADK64084
ID ADK64084 standard; protein; 706 AA.

XX AC ADK64084;

XX DT 06-MAY-2004 (first entry)

XX DE Disease treating protein complex-derived protein #1333.

XX KW protein complex; drug target; diagnosis.

XX OS Unidentified.

XX PN EP1338608-A2.

XX PD 27-AUG-2003.

XX PF 20-DEC-2002; 2002EP-00102902.

XX PR 20-DEC-2001; 2001EP-00130253.

XX PA (CELL-) CELLZONE AG.

XX PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;

XX PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;

XX PI Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.

XX DR N-PSDB; ADK64085.

XX PT New proteins and protein complexes from eukaryotes, useful as targets in
XX PT drug screening, or in diagnosing or screening for the presence of a
XX PT disease or disorder, or a predisposition for developing a disease or
XX PT disorder in a subject.

XX PS Disclosure; SEQ ID NO 2665; 13pp; English.

XX CC The invention relates to novel protein complexes comprising a first and a
XX CC second protein, or its derivative, fragment, homologue or variant. The
XX CC proteins are selected from given protein complexes, which are not defined
XX CC in the specification. The variants are encoded by nucleic acids that
XX CC hybridize to the nucleic acids encoding the proteins under low stringency
XX CC conditions. The protein complexes are useful as targets for an active
XX CC agent of a pharmaceutical. These protein complexes are particularly
XX CC useful as drug targets for the treatment or preventing of a disease or
XX CC disorder. The complexes and methods above are useful in diagnosing or
XX CC screening for the presence of a disease or disorder or a predisposition
XX CC for developing a disease or disorder in a subject. These are also useful
XX CC in screening for a drug for treatment or prevention of a disease or
XX CC disorder. The molecule that modulates the amount, activity or protein
XX CC components of the complex is useful for the manufacture of a medicament
XX CC for the treatment or prevention of a disease or disorder. This sequence
XX CC corresponds to a protein of the invention. (Note: the sequence data for
XX CC this patent did not form part of the printed specification but was
XX CC obtained from the EPO in electronic format).

XX SQ Sequence 706 AA;

Query Match 69.4%; Score 43; DB 7; Length 706;

Best Local Similarity 81.8%; Pred. No. 4e+02;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LEENKELENAL 12
Db 22 LEENKEENSL 32
||||| |.:

RESULT 21

ADL70610

XX ID ADL70610 standard; peptide; 17 AA.

XX AC ADL70610;

XX DT 20-MAY-2004 (first entry)

XX DE Human thrombospondin-1 epitope, useful for cancer diagnosis.

XX KW Human; thrombospondin-1; epitope; cancer; diagnosis.

XX OS Homo sapiens.

XX PN WO2004018995-A2.

XX PD 04-MAR-2004.

XX PF 20-AUG-2003; 2003WO-US026023.

XX PR 23-AUG-2002; 2002US-0405494P.

XX PR 21-APR-2003; 2003US-00419462.

XX PA (WILL/) WILLIAMS K J.

XX PI Williams KJ;

XX WPI; 2004-226901/21.

XX PT New purified thrombospondin fragment extracted from a body fluid, useful
XX PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
XX PT or leukemia or as calibrators, indicators, immunogens and analytes.

XX PS Claim 10; SEQ ID NO 9; 76pp; English.

XX CC The present sequence is that of an epitope of human thrombospondin-1
XX CC (TSP) ADL70639. The epitope is found in TSP and also in a TSP fragment
XX CC present in plasma. The invention relates to TSP fragments (80-110, 40-60
XX CC or 20-35 kDa) found in plasma, and their use in clinical assays for
XX CC cancer and for generation of antibodies and other binding agents. A
XX CC method that distinguishes TSP from a TSP fragment or portion involves:
XX CC (1) using an epitope shared by TSP and the TSP fragment or portion as a
XX CC target for a binding molecule, e.g. an antibody, to obtain a quantitation
XX CC of TSP plus TSP fragment or portion; (2) using an epitope present in TSP
XX CC but not in the fragment or portion to obtain a quantitation of TSP only;
XX CC and (3) using the difference between (1) and (2) as a quantitation of the
XX CC amount of TSP fragment or portion. The present epitope can be used in
XX CC step (1). Detection or quantification of the TSP fragment or portion is
XX CC performed in order to detect the presence, or monitor the course, of a
XX CC disease or condition selected from cancer, renal failure, renal disease,
XX CC atopic dermatitis, vasculitis, acute vasculitis, renal allograft, asthma,
XX CC diabetes mellitus, myocardial infarction, liver disease, splenectomy,
XX CC dermatomyositis, polyarteritis nodosa, systemic lupus erythematosus,
XX CC lupus erythematosus, Kawasaki syndrome, non-specific vasculitis, juvenile
XX CC rheumatoid arthritis, rheumatoid arthritis, vasculitis syndrome, Henoch-
XX CC Scheinlein purpura, thrombocytopenic purpura, purpura, an inflammatory
XX CC condition, a condition associated with clotting, a condition associated
XX CC with platelet activation, a condition associated with consumption of
XX CC platelet activation, a condition associated with consumption of
XX CC platelets, heparin-induced thrombocytopenia, disseminated intravascular
XX CC coagulation, intravascular coagulation, extravascular coagulation, a
XX CC condition associated with endothelial activation, a condition associated
XX CC with production and/or release of thrombospondin and/or a thrombospondin
XX CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
XX CC reaction, an aspartame reaction, atopic dermatitis, eczema,
XX CC hypersensitivity, scleroderma, conditions associated with plugging of
XX CC vessels, a condition associated with a cryofibrinogen, a condition
XX CC associated with a cryoglobulin, and a condition associated with an anti-
XX CC cardiolipin antibody. The cancer is selected from adenoma,
XX CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
XX CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,

CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system,
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.
 XX
 SQ Sequence 17 AA;

Query Match 67.7%; Score 42; DB 8; Length 17;
 Best Local Similarity 61.5%; Pred. No. 11;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELNALK 13
 : ||||| :
 DB 3 VTEENKELNELR 15

RESULT 22

ADL70611

ID ADL70611 standard; peptide; 18 AA.

AC ADL70611;

DT 20-MAY-2004 (first entry)

XX Human thrombospondin-1 epitope, useful for cancer diagnosis.

DE Human; thrombospondin-1; epitope; cancer; diagnosis.

KW Human; thrombospondin-1; epitope; cancer; diagnosis.

XX Homo sapiens.

OS WO2004018995-A2.
 PN 04-MAR-2004.

XX 20-AUG-2003; 2003WO-US026023.

XX 23-AUG-2002; 2002US-0405494P.

PR 21-APR-2003; 2003US-00419462.

XX (WILL/) WILLIAMS K J.

PA Williams KJ;

PI WPI; 2004-226901/21.

XX New purified thrombospondin fragment extracted from a body fluid, useful

PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
 PT or leukemia or as calibrators, indicators, immunogens and analytes.

XX Claim 10; SEQ ID NO 10; 76pp; English.

XX The present sequence is that of an epitope of human thrombospondin-1
 CC (TSP) ADL70639. The epitope is found in TSP and also in a TSP fragment
 CC present in plasma. The invention relates to TSP fragments (80-110, 40-60
 CC or 20-35 kDa) found in plasma, and their use in clinical assays for
 CC cancer and for generation of antibodies and other binding agents. A
 CC method that distinguishes TSP from a TSP fragment or portion involves:
 CC (1) using an epitope shared by TSP and the TSP fragment or portion as a
 CC target for a binding molecule, e.g. an antibody, to obtain a quantitation
 CC of TSP plus TSP fragment or portion; (2) using an epitope present in TSP
 CC but not in the fragment or portion to obtain a quantitation of TSP only;
 CC and (3) using the difference between (1) and (2) as a quantitation of the
 CC amount of TSP fragment or portion. The present epitope can be used in
 CC step (1). Detection or quantification of the TSP fragment or portion is
 CC performed in order to detect the presence, or monitor the course, of a
 CC disease or condition selected from cancer, renal failure, renal disease,
 CC atopic dermatitis, vasculitis, acute vasculitis, renal allograft, asthma,
 CC diabetes mellitus, myocardial infarction, liver disease, splenectomy,

CC dermatomyositis, polyarteritis nodosa, systemic lupus erythematosus,
 CC lupus erythematosus, Kawasaki syndrome, non-specific vasculitis, juvenile
 CC rheumatoid arthritis, rheumatoid arthritis, vasculitis syndrome, Henoch-
 CC Schoenlein purpura, thrombocytopenic purpura, purpura, an inflammatory
 CC condition, a condition associated with clotting, a condition associated
 CC with platelet activation, a condition associated with consumption of
 CC platelets, heparin-induced thrombocytopenia, disseminated intravascular
 CC coagulation, intravascular coagulation, extravascular coagulation, a
 CC condition associated with endothelial activation, a condition associated
 CC with production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition
 CC associated with a cryoglobulin, and a condition associated with an anti-
 CC cardiophilin antibody. The cancer is selected from adenoma, a
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system, liver or
 CC muscle, bone, a joint, tendon or ligament, digestive system, reproductive
 CC biliary system, pancreas, head, neck, endocrine system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.
 XX

SQ Sequence 18 AA;

Query Match 67.7%; Score 42; DB 8; Length 18;
 Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELNALK 13
 : ||||| :
 DB 4 VTEENKELNELR 16

RESULT 23

AAO04456

ID AAO04456 standard; protein; 52 AA.

AC AAO04456;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 18348.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AA184387.

XX

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 18348; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 52 AA;

Query Match 67.7%; Score 42; DB 4; Length 52;

Best Local Similarity 61.5%; Pred. No. 36;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ILLENKELENALK 13

Db 1 LLERNKELEGSQ 13

RESULT 24

ADL70643
 ID ADL70643 standard; protein; 98 AA.

XX ADL70643;

XX 20-MAY-2004 (first entry)

XX Human thrombospondin-1 procollagen homology domain.

XX Human; thrombospondin-1; epitope; cancer; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 71..98 /note= "Collagen V binding region (N-terminal portion)"

XX WO2004018995-A2.

XX 04-MAR-2004.

XX 20-AUG-2003; 2003WO-US026023.

XX 23-AUG-2003; 2002US-0405494P.

XX 21-APR-2003; 2003US-00419462.

XX (WILL/) WILLIAMS K J.

XX Williams KJ;

XX WPI; 2004-226901/21.

XX New purified thrombospondin fragment extracted from a body fluid, useful
 PT for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
 PT or leukemia or as calibrators, indicators, immunogens and analytes.

XX Disclosure; SEQ ID NO 42; 76pp; English.

XX The present sequence is that of the procollagen homology domain of human
 CC thrombospondin-1 (TSP) ADL70639. It includes the N-terminal portion of
 CC the collagen V binding region; the remainder of this region lies in the
 CC type 1 repeat domain ADL70644. The invention relates to TSP fragments (80

CC -100, 40-55 or 20-35 kDa mol.wt.) found in plasma, and their use in
 CC clinical assays for cancer and for generation of antibodies and other
 CC binding agents. A method that distinguishes TSP from a TSP fragment or
 CC portion involves: (1) using an epitope shared by TSP and the TSP fragment
 CC or portion as a target for a binding molecule, e.g. an antibody, to
 CC obtain a quantitation of TSP plus TSP fragment or portion; (2) using an
 CC epitope present in TSP but not in the fragment or portion to obtain a
 CC quantitation of TSP only; and (3) using the difference between (1) and
 CC (2) as a quantitation of the amount of TSP fragment or portion. Suitable
 CC epitopes are provided ADL70602-ADL70638. Detection or quantification of
 CC the TSP fragment or portion is performed in order to detect the presence,
 CC or monitor the course, of a disease or condition selected from cancer,
 CC renal failure, renal disease, atopic dermatitis, vasculitis, acute
 CC vasculitis, renal allograft, asthma, diabetes mellitus, myocardial
 CC infarction, liver disease, splenectomy, dermatomyositis, polyarthritis
 CC nodosa, systemic lupus erythematosus, lupus erythematosus, Kawasaki
 CC syndrome, non-specific vasculitis, juvenile rheumatoid arthritis,
 CC rheumatoid arthritis, vasculitis syndrome, Henoch-Schoenlein purpura,
 CC thrombocytopenic purpura, purpura, an inflammatory condition, a condition
 CC associated with clotting, a condition associated with platelet
 CC activation, a condition associated with intravascular platelet
 CC activation, a condition associated with consumption of platelets, heparin
 CC -induced thrombocytopenia, disseminated intravascular coagulation,
 CC intravascular coagulation, extravascular coagulation, a condition
 CC associated with endothelial activation, a condition associated with
 CC production and/or release of thrombospondin and/or a thrombospondin
 CC fragment, urticaria, hives, angioedema, a drug reaction, an antibiotic
 CC reaction, an aspartame reaction, atopic dermatitis, eczema,
 CC hypersensitivity, scleroderma, conditions associated with plugging of
 CC vessels, a condition associated with a cryofibrinogen, a condition
 CC associated with a cryoglobulin, and a condition associated with an anti-
 CC adenocarcinoma antibody. The cancer is selected from adenoma,
 CC adenocarcinoma, carcinoma, lymphoma, leukaemia, solid cancer, liquid
 CC cancer, metastatic cancer, pre-metastatic cancer, non-metastatic cancer,
 CC a cancer with vascular invasion, internal cancer, skin cancer, cancer of
 CC the respiratory system, circulatory system, musculoskeletal system, of
 CC muscle, bone, a joint, tendon or ligament, digestive system, liver or
 CC biliary system, pancreas, head, neck, endocrine system, reproductive
 CC system (male or female), genitourinary system, kidney, urinary tract,
 CC sensory system, nervous system, lymphoid organ, blood, a gland, mammary
 CC gland, prostate gland, endometrial tissue, mesodermal tissue, ectodermal
 CC tissue, endodermal tissue, a teratoma, a poorly-differentiated cancer, a
 CC well-differentiated cancer or a moderately differentiated cancer.

XX Sequence 98 AA;

Query Match 67.7%; Score 42; DB 8; Length 98;

Best Local Similarity 61.5%; Pred. No. 70;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILLENKELENALK 13

Db 21 VTEENKELANEUR 33

RESULT 25

ADQ39359

ID ADQ39359 standard; protein; 432 AA.

XX ADQ39359;

XX 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1022.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 XX cardiac; gene therapy; human.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.


```
XX 22-DEC-2003; 2003WO-US040978.
PF
XX
XX 20-DEC-2002; 2002US-0434778P.
PR
XX 10-MAR-2003; 2003US-0453135P.
PR
XX 30-APR-2003; 2003US-0466412P.
PR
XX 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.
PA
XX Cargill M, Devlin JJ, Iakubova O;
PI
XX WPI; 2004-533949/51.
DR
XX N-PSDB; ADQ38531.
DR
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
PT
XX Claim 10; SEQ ID NO 1022; 145pp; English.
PS
XX
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
XX Sequence 432 AA;
SQ
Query Match 67.7%; Score 42; DB 8; Length 432;
Best Local Similarity 61.5%; Pred. No. 3.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
: ||||| |
DB 301 VTEENKELANELR 313
RESULT 26
ADQ39357
ID ADQ39357 standard; protein; 432 AA.
XX
XX ADQ39357;
AC
XX 18-NOV-2004 (first entry)
DT
XX Human myocardial infarction-associated gene derived protein, SEQ ID 1020.
DE
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX
```

```
OS Homo sapiens.
XX
XX WO2004058052-A2.
XX
XX 15-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US040978.
XX
XX 20-DEC-2002; 2002US-0434778P.
PR
XX 10-MAR-2003; 2003US-0453135P.
PR
XX 30-APR-2003; 2003US-0466412P.
PR
XX 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.
PA
XX Cargill M, Devlin JJ, Iakubova O;
PI
XX WPI; 2004-533949/51.
DR
XX N-PSDB; ADQ38529.
DR
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
PT
XX Claim 10; SEQ ID NO 1020; 145pp; English.
PS
XX
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
XX Sequence 432 AA;
SQ
Query Match 67.7%; Score 42; DB 8; Length 432;
Best Local Similarity 61.5%; Pred. No. 3.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
: ||||| |
DB 301 VTEENKELANELR 313
RESULT 27
AAU02916
ID AAU02916 standard; protein; 459 AA.
XX
XX AAU02916;
AC
XX 12-SEP-2001 (first entry)
DT
XX
```

DE Angiotensin converting enzyme (ACEV) splice variant protein #16.
XX
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroiditic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX
OS Homo sapiens.
XX
PN WO200136632-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-IL000766.
XX
PR 17-NOV-1999; 99IL-00132978.
PR 10-DEC-1999; 99IL-00133455.
XX
PA (COMP-) COMFUGEN LTD.
XX
PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX
DR WPI; 2001-336004/35.
DR N-PSDB; AAS06016.
XX
PT Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
XX
PS Claim 4; Fig 16; 519pp; English.
XX
CC The sequence represents an angiotensin converting enzyme splice variant
CC (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen P53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroiditic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
CC abnormality such as deep vein thrombosis
XX
SQ Sequence 459 AA;
Query Match 67.7%; Score 42; DB 4; Length 459;
Best Local Similarity 61.5%; Pred. No. 3.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
RESULT 28
ABG74686
ID ABG74686 standard; protein; 465 AA.
XX
AC ABG74686;
XX
DT 10-MAY-2003 (first entry)
XX

DE Human CGDD protein 7727886CD1 SEQ ID 12.
XX
KW Human; cell growth, differentiation and death; CGDD; cardiact; cancer;
KW cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
KW neurological disorder.
XX
OS Homo sapiens.
XX
PN WO2003014322-A2.
XX
PD 20-FEB-2003.
XX
PF 08-AUG-2002; 2002WO-US025465.
XX
PR 08-AUG-2001; 2001US-0311017P.
PR 17-AUG-2001; 2001US-0313070P.
PR 17-AUG-2001; 2001US-0313071P.
PR 24-AUG-2001; 2001US-0314678P.
PR 31-AUG-2001; 2001US-0316692P.
PR 07-SEP-2001; 2001US-0317913P.
PR 14-SEP-2001; 2001US-0322182P.
PR 07-DEC-2001; 2001US-0340747P.
PR 20-DEC-2001; 2001US-0342761P.
PR 29-MAR-2002; 2002US-0369129P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;
PI Duggan BM, Elliott VS, Emerling BM, Forsythe IJ, Gietzen KJ;
PI Corvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
PI Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
PI Yue H;
XX
DR WPI; 2003-268197/26.
DR N-PSDB; ABQ77425.
XX
PT New polypeptide associated with cell growth, differentiation and death,
PT useful for preparing a composition for diagnosing or treating a disease
PT e.g., cardiovascular or neurological disorders or cancer.
XX
PS Claim 67; Page 215-216; 267pp; English.
XX
CC This invention describes a novel polypeptide associated with cell growth,
CC differentiation and death (CGDD) which has cardiact, cytostatic and
CC neuroprotective activity. The polypeptides of the invention are useful
CC for preparing a composition for diagnosing, or treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional CGDD e.g., cardiovascular or neurological disorders or cancer.
CC The polynucleotides of the invention can be used for gene therapy.
CC ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
CC ABQ77414-ABQ77440
XX
SQ Sequence 465 AA;
Query Match 67.7%; Score 42; DB 6; Length 465;
Best Local Similarity 61.5%; Pred. No. 3.7e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 38 LLERNKELEGSIQ 50
RESULT 29
AAB43602
ID AAB43602 standard; protein; 466 AA.
XX
AC AAB43602;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1047.

XX	Human; cancer associated gene; cancer antigen; detection; cancer;
KW	diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW	antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW	dermatological; neuroprotective; thrombolytic; coagulant; neotropic;
KW	vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW	allergic reaction; graft versus host disease; organ rejection;
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
KW	neurological disease; drug screening.
OS	Homo sapiens.
XX	WO200055350-A1.
XX	21-SEP-2000.
XX	08-MAR-2000; 2000WO-US005892.
XX	12-MAR-1999; 99US-0124270P.
PR	(HUMA-) HUMAN GENOME SCI INC.
PA	Rosen CA, Ruben SM;
XX	PI
XX	XX
DR	WPI; 2000-587533/55.
DR	N-PSDB; AAC77811.
XX	Novel isolated nucleic acids comprising sequences encoding peptides
PT	useful for treating or diagnosing e.g. cancer.
PT	Claim 11; Page 1636-1638; 2352pp; English.
XX	AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC	AAB43398 to AAB44239. The proteins can have activities based on the
CC	tissues and cells the genes are expressed in. Example of activities
CC	include: cytostatic; proliferative; vulnery; immunomodulator;
CC	antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC	antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC	dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC	neotropic; vasotropic; antiproliferative and angiogenic. The
CC	polynucleotides and polypeptides can be used for preventing, treating or
CC	ameliorating medical conditions and diagnosing pathological conditions.
CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC	the present invention may be used to treat immune disorders by activating
CC	or inhibiting the proliferation, differentiation or mobilisation of
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune
CC	disorders, allergic reactions, graft versus host disease and organ
CC	rejection, modulate haemostatic or thrombolytic activity, modulate
CC	inflammation, cancers, cardiovascular disorders, neurological disease and
CC	bacterial or viral infections. The peptides, nucleotides, antibodies,
CC	agonists and antagonists may be also be used in drug screens. AAC78449 of
CC	AAC78457 and AAB44240 represent sequences used in the exemplification of
CC	the present invention
XX	Sequence 466 AA;
XX	Sequence 466 AA;
XX	Query Match 67.7%; Score 42; DB 3; Length 466;
XX	Best Local Similarity 61.5%; Pred. No. 3.7e+02;
XX	Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy	1 ILEENKELENALK 13
Db	:
	354 VTEENKELANEUR 366
RESULT 30	
ADY49758	
ID ADY49758	standard; protein; 534 AA.
XX	
XX	ADY49758;
XX	

XX 17-JUN-2004 (first entry)
DT TSF polypeptide.
DE adenovirus vector; anti-neoplastic; TSF polypeptide; cancer.
KW Homo sapiens.
XX CN1401387-A.
OS 12-MAR-2003.
PN 21-AUG-2002; 2002CN-00129408.
PD 21-AUG-2002; 2002CN-00129408.
XX (TAID-) TAIDA LIFE SCI TECH RES CENT HEMATOLOGY.
PI Han Z, Liu P;
XX WPI; 2003-469302/45.
DR N-PSDB; ADN02475.
XX Tumor suppressing polypeptide TSF and gene therapy vector composition.
PT Claim 2; SEQ ID NO 1; 13pp; Chinese.
PS The present invention relates to a novel recombinant adenovirus vector
CC mediated anti-neoplastic composition is prepared through cloning the cDNA
CC sequence from the human peripheral blood cell by specific primer and
CC reverse transcription-polymerase chain reaction (RT-PCR) method for
CC coding TSF polypeptide, construction in human embryonic kidney cell 293
CC by AdEasy system, and packaging and expressing the recombinant adenovirus
CC vector of TSF. It can suppress the growth and transfer of cancer. The
CC present sequence represents the TSF polypeptide.
XX Sequence 548 AA;
SQ Query Match 67.7%; Score 42; DB 7; Length 548;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB : ||||| | :
301 VTEENKELANELR 313
RESULT 34
AAU02914
ID AAU02914 standard; protein; 555 AA.
XX AAU02914;
AC 12-SEP-2001 (first entry)
DT Angiotensin converting enzyme (ACEV) splice variant protein #14.
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX Homo sapiens.
OS WO200136632-A2.
PN
XX

PD 25-MAY-2001.
XX 17-NOV-2000; 2000WO-IL000766.
PF 17-NOV-1999; 99IL-00132978.
XX 10-DEC-1999; 99IL-00133455.
PR (COMP-) COMPUGEN LTD.
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
PI WPI; 2001-336004/35.
XX N-PSDB; AAS06014.
DR Novel alternative splicing variants e.g. variant of angiotensin
XX converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies.
PT Claim 4; Fig 14; 519pp; English.
PS The sequence represents an angiotensin converting enzyme splice variant
XX (ACEV) polypeptide. The polypeptides of the invention include variants of
CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
CC polypeptide receptor 2. The polypeptides and their associated nucleic
CC acids are useful for identification of variant sequences and detection of
CC candidate compounds capable of binding the molecules. The sequences of
CC the invention can be used in the treatment and diagnosis of various
CC disorders including cardiovascular diseases such as arteriosclerosis,
CC myocardial infarction and coronary arterial thrombosis, renal diseases
CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
CC immune disorders such as immune complex nephritis, multiple sclerosis,
CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
CC as asbestosis and vascular pathologies involving an endothelial
XX abnormality such as deep vein thrombosis
SQ Sequence 555 AA;
Query Match 67.7%; Score 42; DB 4; Length 555;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB : ||||| | :
301 VTEENKELANELR 313
RESULT 35
ADY08679
ID ADY08679 standard; protein; 625 AA.
XX AC ADY08679;
AC 21-APR-2005 (first entry)
DT Plant full length insert polypeptide seqid 64494.
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX Unidentified.
OS US2004034888-A1.
XX 19-FEB-2004.
PN
PD
XX

PF 28-APR-2003; 2003US-004251114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI; 2004-180133/17.
 DR
 XX
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 64494; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 625 AA;
 Query Match 67.7%; Score 42; DB 8; Length 625;
 Best Local Similarity 61.5%; Pred. No.: 5e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ILEENKELENALK 13
 : ||||| :
 Db 432 VLEEKAELENRMK 444
 RESULT 36
 AAU02913
 ID AAU02913 standard; protein; 731 AA.
 XX
 AC AAU02913;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Angiotensin converting enzyme (ACEV) splice variant protein #13.
 XX
 KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
 KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.

XX Homo sapiens.
 XX WO200136632-A2.
 XX 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000WO-IL000766.
 XX
 PR 17-NOV-1999; 99IL-00132978.
 PR 10-DEC-1999; 99IL-00133455.
 XX
 PA (COMP-) COMPUGEN LTD.
 XX
 PI Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 XX WPI; 2001-336004/35.
 DR N-PSDB; AAS06013.
 XX
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies.
 XX
 PS Claim 4; Fig 13; 51pp; English.
 XX
 CC The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding to the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis
 XX
 SQ Sequence 731 AA;
 Query Match 67.7%; Score 42; DB 4; Length 731;
 Best Local Similarity 61.5%; Pred. No.: 6e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ILEENKELENALK 13
 : ||||| :
 Db 301 VTEENKELANELR 313
 RESULT 37
 AAB00042
 ID AAB00042 standard; protein; 1152 AA.
 XX
 AC AAB00042;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human thrombospondon-1 (TSP-1).
 XX
 KW TSP-1; TSP-2; COMP; cartilage oligomeric matrix protein; thrombospondin;
 KW angiogenesis; tumour; treatment; cancer; arthritis; psoriasis;
 KW diabetic retinopathy; corneal graft rejection; glaucoma.
 XX
 OS Homo sapiens.
 XX
 XX Location/Qualifiers
 FH Key 361..416
 FT Region /label= Type 1 repeat region
 FT Region 417..473

FT /label= Type 1 repeat region
FT 474..530
FT /label= Type 1 repeat region
XX
PN WO200044908-A2.
PD
XX 03-AUG-2000.
XX
PF 01-FEB-2000; 2000WO-US02482.
XX
PR 01-FEB-1999; 99US-0118053P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2000-514823/46.
DR
XX Nucleic acids encoding chimeric proteins such as cartilage oligomeric
PT matrix protein (COMP)/thrombospondins (TSP)-1 and 2, useful for
PT inhibiting angiogenesis and treating diseases such as cancer.
XX
XX Disclosure; Fig 1; 40pp; English.
XX
XX New nucleic acids are described which encode a protein comprising the
CC second and third type 1 repeats of human TSP (thrombospondin)-1, but not
CC the TGF (transforming growth factor)-beta activation region of human TSP-
CC 1. The nucleic acid of TSP (thrombospondin)-1 containing the second and
CC third type-1 repeats and the COMP (cartilage oligomeric matrix protein)
CC assembly sequence (COMP/TSP-1) was produced by PCR (polymerase chain
CC reaction). Expression of COMP/TSP-1 caused inhibition of the growth of
CC tumours in mice models. Thus the nucleic acids and proteins may be useful
CC for treating angiogenesis related diseases such as cancer (by reducing
CC the rate of growth and size of tumours), arthritis, psoriasis, diabetic
CC retinopathy, corneal graft rejection, and glaucoma. They may also be used
CC for treating human immunodeficiency virus (HIV) infection. Anti-
CC angiogenic therapy has little toxicity, does not require the therapeutic
CC agent to enter tumour cells or cross the blood-brain barrier, controls
CC tumour growth independently of growth of tumour cell heterogeneity, and
CC does not induce drug resistance
XX
XX Sequence 1152 AA;
SQ

Query Match 67.7%; Score 42; DB 3; Length 1152;
Best Local Similarity 61.5%; Pred. NO. 9.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILLENKELENALK 13
 : ||||| | | :
DB 283 VTEENKELANEUR 295

RESULT 38
AAU74771
ID AAU74771 standard; protein; 1152 AA.
XX
AC AAU74771;
XX
DT 09-APR-2002 (first entry)
XX
DE Human thrombospondin-1 (TSP-1).
XX
XX Thrombospondin-1; TSP-1; cytostatic; angiogenesis; vasotropic; vulnary;
KW neovascularisation; cell proliferation inhibitor; cancer; solid tumour;
KW haemangioma; acoustic neuromas; neurofibroma; trachoma;
KW pyogenic granulomas; rheumatoid arthritis; ocular angiogenic disease;
KW retinopathy; psoriasis; macular degeneration; corneal graft rejection;
KW neovascular glaucoma; retrolental fibroplasia; reboosis; angiofibroma;
KW Osler-Webber syndrome; myocardial angiogenesis; haemophilic joints;
KW plaque neovascularisation; telangiectasia; wound granulation; apoptosis.
XX
OS Homo sapiens.
XX

PH Key Location/Qualifiers
FT Region 263..360
FT /label= Procollagen_homology_region
FT 303..309
FT /label= Procollagen_homology_domain
FT /label= "Required in inhibition of angiogenesis"
FT 361..530
FT /label= Type 1 repeat domain
FT /label= "This region contains 3 type 1 repeats, from
FT residues 361-416, residues 417-473 and residues 474-530"
FT 364..370
FT /label= Heparin_binding_domain
FT 413..415
FT /label= RFX motif
FT /label= "Necessary and sufficient for activation of
FT transforming growth factor beta (TGF beta)"
FT 418..423
FT /label= TGF-beta and fibronectin binding_domain
FT /label= "Transforming growth factor"
FT 420..426
FT /label= Heparin_binding_domain
FT 429..434
FT /label= Cell_binding_domain
FT 477..483
FT /label= Heparin_binding_domain
FT 481..499
FT /label= Anti-angiogenesis_domain
FT 486..491
FT /label= Cell_binding_domain
FT 531..673
FT /label= Type 2 repeat domain
FT /label= "This region contains 3 type 2 repeats, from
FT residues 531-571, residues 572-629 and residues 630-673"
FT 570..601
FT /label= Calcium_binding_domain
FT 698..925
FT /label= Type 3 repeat domain
FT /label= "This region contains 7 type 3 repeats, from
FT residues 698-733, residues 734-756, residues 757-782, and
FT residues 793-815, residues 816-853, residues 854-889, and
FT residues 890-925"
XX
XX WO200191781-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US017250.
XX
XX 26-MAY-2000; 2000US-0207994P.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
XX Lawler JW;
XX
XX WPI; 2002-106273/14.
XX
XX Composition useful for treatment of cancer comprises cDNA encoding amino
PT acids of human thrombospondin-1 or its conservative variant and a
PT carrier.
XX
XX Disclosure; Fig 7; 54pp; English.
XX
XX The invention describes a composition comprising cDNA encoding fragments
CC of human thrombospondin-1 (TSP -1), a type 1 repeat polypeptide and
CC potent inhibitor of tumour growth and angiogenesis. The composition is
CC useful for killing cancerous cells (preferably tumour); for reducing
CC volume or inhibiting growth of a tumour (inhibiting neovascularisation in
CC the tumour); for decreasing proliferation of tumour cells; in the
CC treatment of diseases and conditions associated with angiogenic activity
CC or misregulated growth and angiogenesis-mediated diseases such as cancer,
CC solid tumour, tumour metastasis, benign tumour, (e.g. haemangioma,
CC acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas),
CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases (e.g.

CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,
CC corneal graft rejection, neovascular glaucoma, retrolental fibroplasias,
CC rebeosis), Osler-Weber syndrome, myocardial angiogenesis,
CC telangiectasia, plaque neovascularisation, haemophilic joints,
CC angiofibroma or wound granulation. The composition induces apoptosis and
CC inhibits neovascularisation in the tumour cells. This amino acid sequence
CC represents human thrombospondin-1 (TSP-1), on which the recombinant
XX proteins of the invention are based
XX
SQ Sequence 1152 AA;

Query Match 67.7%; Score 42; DB 5; Length 1152;
Best Local Similarity 61.5%; Pred. NO. 9.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
DB 283 VTEENKELANELR 295

RESULT 39
ABB82285
ID ABB82285 standard; protein; 1152 AA.
XX
AC ABB82285;
XX
DT 08-JAN-2003 (first entry)
XX

DE Human thrombospondin-1 (TSP-1) protein sequence.

XX TSP-1; carrier protein; chimeric; multimerization domain; spacer;
KW therapeutic; gene therapy; human; thrombospondin-1.
XX
XX Homo sapiens.

Key	Location/Qualifiers
Region	263..360
Peptide	/note= "Procollagen homology region"
Peptide	361..416
Peptide	/note= "type 1 repeat"
Peptide	417..473
Peptide	/note= "type 1 repeat"
Peptide	474..530
Peptide	/note= "type 1 repeat"
Peptide	531..571
Peptide	/note= "type 2 repeat"
Peptide	572..629
Peptide	/note= "type 2 repeat"
Peptide	630..673
Peptide	/note= "type 2 repeat"
Peptide	674..697
Peptide	/note= "type 2 repeat"
Peptide	698..733
Peptide	/note= "type 3 repeat"
Peptide	734..756
Peptide	/note= "type 3 repeat"
Peptide	757..792
Peptide	/note= "type 3 repeat"
Peptide	793..815
Peptide	/note= "type 3 repeat"
Peptide	816..853
Peptide	/note= "type 3 repeat"
Peptide	854..889
Peptide	/note= "type 3 repeat"
Peptide	890..925
Peptide	/note= "type 3 repeat"

XX WO200270725-A1.

XX 12-SEP-2002.

XX 06-MAR-2002; 2002WO-US006882.

XX

PR 06-MAR-2001; 2001US-0273573P.

XX (ABDE/) ABDELOUAHED M.

XX Abdelouahed M;

XX WPI; 2002-750463/81.

XX New carrier chimeric protein having a multimerization domain and a spacer
PT to which at least one drug is attached, useful as therapeutic agent, or in
PT gene therapy.

XX Example 1; Fig 2; 32pp; English.

XX The invention relates to a carrier chimeric protein, having a
CC multimerization domain and a spacer to which at least one drug is
CC attached. The carrier chimeric protein comprises the formulae (F1): S1-A-
CC S2-B and (F2): T-S1-A-S2-B, where T = S1 = S2 = spacer; A = B; where T =
CC specific amino acid domain to a specific site of action; S1 = spacer and
CC is present or absent; S2 = spacer; A = multimerization domain of a
CC tetrameric or pentameric protein, or analogous products; and B = protein-
CC or amino acid-drug. When S1 is present, S1 and S2 are identical or
CC different. The carrier chimeric protein is useful as therapeutic agent or
CC in gene therapy. The present sequence represents the amino acid sequence
CC of human thrombospondin-1 (TSP-1), the fragments of which can be used in
CC the construction of the carrier chimeric protein
XX
SQ Sequence 1152 AA;

Query Match 67.7%; Score 42; DB 5; Length 1152;

Best Local Similarity 61.5%; Pred. NO. 9.7e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 283 VTEENKELANELR 295

RESULT 40

AA874450

ID AA874450 standard; protein; 1170 AA.

XX AA874450;

XX 06-JUN-2001 (first entry)

XX Human variant thrombospondin 1.

XX Human; variant thrombospondin 1; variant thrombospondin 4; SNP;

KW polymorphism; vascular disease; coronary artery disease; forensics;

KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;

XX pulmonary embolism; paternity test.

XX Homo sapiens.

XX WO200118250-A2.

XX 15-MAR-2001.

XX 07-SEP-2000; 2000WO-US024503.

XX 10-SEP-1999; 99US-0153357P.

XX 26-JUL-2000; 2000US-0220947P.

XX 16-AUG-2000; 2000US-0225724P.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;

XX WPI; 2001-226749/23.

XX N-PSDB; AAF95238.

XX

XX Nucleic acids comprising single nucleotide polymorphisms, useful in
PT applications such as forensics, paternity testing, medicine, genetic
PT analysis and phenotype correlations to diseases such as diabetes and
PT atherosclerosis.
XX
XX
PS Claim 20; Fig 1; 242pp; English.
XX
CC The present invention provides a method of diagnosing a vascular disease
CC in an individual, involving determining the sequence at various
CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC genes. The sequences at a number of polymorphic sites are also provided
CC in the specification. In particular, the method can be used in the
CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC disease, stroke, peripheral vascular diseases, venous thromboembolism and
CC pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC useful in forensics, paternity testing, genetic analysis and phenotype
CC correlations to diseases. The present sequence is the human variant
CC thrombospondin 1 protein
XX
XX Sequence 1170 AA;
SQ
Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
RESULT 41
ID AAB90800 standard; protein; 1170 AA.
XX
AC AAB90800;
DT 15-JUN-2001 (first entry)
XX
DE Human shear stress-response protein; vascular disease; arteriosclerosis.
KW Human; shear stress-response protein; vascular disease; arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO200125427-A1.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-JP006840.
XX
PR 01-OCT-1999; 99JP-00280976.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (NOJIT/) NOJIMA H.
XX
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
PI Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
XX WPI; 2001-266308/27.
DR N-PSDB; AA02923.
XX
XX DNA sequences, proteins encoded by them and antibodies against them
PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis.
XX
XX Claim 60; Page 515-521; 678pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension

XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
RESULT 42
ID AAE25030 standard; protein; 1170 AA.
XX
AC AAE25030;
DT 30-OCT-2002 (first entry)
XX
DE Human thrombospondin (TSP)-1 protein.
XX
KW Human; cardiovascular disease; CVD; diagnosis; thrombospondin; TSP;
KW anticoagulant therapy; atherosclerosis; coronary artery disease; MI;
KW myocardial infarction; venous thromboembolism; pulmonary embolism;
KW peripheral vascular disease; congestive heart failure; gene therapy;
KW stroke; TSP-1; single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT Misc-difference 700
FT /note= "Asn at this position is changed to Ser due to
FT single nucleotide polymorphism (SNP)"
XX
PN WO200239122-A2.
XX
PD 16-MAY-2002.
XX
PF 09-NOV-2001; 2001WO-US047406.
XX
PR 09-NOV-2000; 2000US-0248185P.
PR 22-DEC-2000; 2000US-0257417P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy J;
XX
XX WPI; 2002-519324/55.
DR N-PSDB; AAD40610.
XX
XX Diagnosing and monitoring a cardiovascular disease in a patient, by
PT assessing the level of thrombospondin protein and/or gene expression in a
PT biological sample.
XX
XX Example 1; Page 97-100; 108pp; English.
XX
XX The invention relates to a method of diagnosing and monitoring a
CC cardiovascular disease (CVD) in a patient. The method involves assessing
CC the level of thrombospondin (TSP) protein and/or gene expression in a
CC biological sample. The invention also provides a method useful for
CC assessing the efficacy of a compound or anticoagulant therapy for
CC inhibiting CVD in a patient. CVD include atherosclerosis, coronary artery
CC disease, myocardial infarction (MI), stroke, peripheral vascular disease,
CC congestive heart failure, venous thromboembolism and pulmonary embolism.
CC The invention is also used in gene therapy. The present sequence is human
CC TSP-1 protein
XX
XX Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
RESULT 43
ID AAU75315 standard; protein; 1170 AA.
XX AAU75315;
AC AAU75315;
XX 23-APR-2002 (first entry)
DT Human thrombospondin 1, TSP-1, protein.
XX Human; Erg; transcription factor; tumour necrosis factor-alpha;
XX TNF-alpha; interleukin-1 (IL-1); SPARC; thrombospondin; TSP; rho-A;
KW intracellular adhesion molecule-2; ICAM-2; von Willebrand factor; vWF;
KW angiogenesis; inflammatory disorder; developmental disorder; wound;
KW menstrual disorder; cancer; rheumatoid arthritis; atherosclerosis;
KW diabetic retinopathy; restenosis; osteoporosis; cataract;
KW diabetes mellitus; glomerulonephritis; inflammatory glomerular disease;
KW vasculitis; retinopathy; liver fibrosis; haemoglobinopathy;
KW Crohn's disease.
XX Homo sapiens.
OS WO200189122-A2.
XX 22-NOV-2001.
XX 15-MAY-2001; 2001WO-GB002140.
XX 16-MAY-2000; 2000US-0204493P.
XX (GLAX) GLAXO GROUP LTD.
XX Mclaughlin F, Randi AM;
XX WPI; 2002-089855/12.
XX New modulator of a member of the Ets family of transcription factors,
XX Erg, which is not tumor necrosis factor-alpha or interleukin-1, for
XX treating cancer, rheumatoid arthritis, atherosclerosis, restenosis, and
XX osteoporosis.
XX Disclosure; Page 55-59; 75pp; English.
XX The invention relates to an Erg (a member of the Ets family of
XX transcription factors) modulator for use in medicine, where the modulator
XX is not tumour necrosis factor-alpha (TNF-alpha) or interleukin-1 (IL-1).
XX Also included are a method of screening for a drug or drug candidate,
XX involving determining whether or not a group being screened modulates the
XX effect of Erg on secreted protein, acidic, cysteine-rich (Osteonectin/BM-
XX 40) (SPARC), thrombospondin (TSP), rho-A (a Ras superfamily member),
XX intracellular adhesion molecule-2 (ICAM-2) and/or von Willebrand factor
XX (vWF) RNA or polypeptide expression, or expression of heterologous RNA or
XX a heterologous polypeptide that is under the control of one or more
XX regulatory sequences of the SPARC, TSP, rho A, ICAM-2 and/or vWF genes,
XX determining whether or not a group being screened modulates the
XX expression of Erg or of Erg RNA in a given expression system, determining
XX whether or not a group being screened modulates binding of Erg to a
XX nucleic acid comprising an Erg binding site, determining whether or not a
XX nucleic acid hybridizes with DNA of an Erg gene or its transcription
XX product or determining whether or not a nucleic acid knocks out or
XX reduces expression of Erg. Erg based probes can be used in the diagnosis
XX of a disorder such as a disorder in which angiogenesis is involved,
XX inflammatory disorder, developmental disorder, wound, menstrual disorder
XX involving cell proliferation, a disorder that can be treated by tissue
XX remodeling, cancer, cancer metastasis, rheumatoid arthritis,
XX atherosclerosis, diabetic retinopathy, restenosis, osteoporosis,
XX cataract, diabetes mellitus, glomerulonephritis, inflammatory glomerular

CC disease, vasculitis, retinopathy, liver fibrosis, haemoglobinopathy or
CC Crohn's disease. The modulator is useful in the preparation of a
CC medicament for treating a disorder which involves aberrant expression of
CC SPARC, TSP, rho A, ICAM-2 or vWF. A group that modulates the effect of
CC Erg or a therapeutically active agent is useful for the preparation of a
CC medicament for treating the above mentioned disorders. The present
CC sequence is the amino acid sequence of human TSP-1
XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
RESULT 44
ABP96780
ID ABP96780 standard; protein; 1170 AA.
XX ABP96780;
AC ABP96780;
XX 05-JUN-2003 (first entry)
DT Human COPD related protein SEQ ID NO:30.
XX Human; chronic obstructive pulmonary disease; COPD; chronic lung disease.
DE Homo sapiens.
OS WO200297127-A2.
XX 05-DEC-2002.
XX 28-MAY-2002; 2002WO-EP005835.
XX 31-MAY-2001; 2001GB-00013266.
XX (FARB) BAYER AG.
XX Oellers N, Gehrman M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX WPI; 2003-140492/13.
XX N-PSDB; ACC46751.
XX Predicting, diagnosing or prognosing chronic lung disease, by detecting a
XX chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX Claim 8; Page 120-125; 214pp; English.
XX The present invention describes a method for predicting, diagnosing or
XX prognosing chronic lung disease by detecting a chronic obstructive
XX pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
XX ACC46777, which encodes the COPD related proteins in ABP96779 to
XX ABP96806). The method is useful for predicting, diagnosing or prognosing
XX chronic lung disease in a biological sample. The COPD genes and proteins
XX encoded by them from the present invention (i) can be used for treating
XX or preventing chronic lung disease in a mammal. (i) can be used in an
XX animal model for determining the efficacy, toxicity, or side effects of
XX treatment with (i), and determining the mechanism of action of (i).
XX ACC46778 to ACC46903 represent COPD related PCR primers and probes used
XX in an example from the present invention
XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13

```
Db          301 VTEENKELANLNR 313

RESULT 45
ABU03474
ID ABU03474 standard; protein; 1170 AA.
XX
AC ABU03474;
XX
XX 21-JAN-2003 (first entry)
XX
DE Angiogenesis-associated human protein sequence #19.
XX
KW Human; angiogenesis-associated transcript; angiogenesis;
KW angiogenesis-associated disease; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200279492-A2.
XX
PD 10-OCT-2002.
XX
PF 14-FEB-2002; 2002WO-US004915.
XX
PR 14-FEB-2001; 2001US-00784356.
XX
PR 22-FEB-2001; 2001US-00781390.
XX
PR 19-APR-2001; 2001US-0285475P.
XX
PR 03-AUG-2001; 2001US-0310025P.
XX
PR 13-NOV-2001; 2001US-0350666P.
XX
PR 29-NOV-2001; 2001US-0334244P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Murray R, Glynn R, Watson SR, Aziz N;
XX
DR WPI; 2003-040681/03.
XX
DR N-PSDB; ABX08757.
XX
XX
PT Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT treating cancer by contacting a sample with a polynucleotide that
PT exhibits changes in expression level as a function of time in tissue
PT undergoing angiogenesis.
XX
PS Example 2; Page 199; 291pp; English.
XX
CC The present invention relates to methods and compositions for detecting
CC an angiogenesis-associated transcript in a cell in a patient. The method
CC involves contacting a biological sample from the patient with a
CC polynucleotide that selectively hybridizes to a sequence at least 80%
CC identical to any of the angiogenesis-associated human polynucleotide
CC sequences given in the specification. These angiogenesis-associated
CC polynucleotide sequences comprise genes that exhibit changes in
CC expression levels as a function of time in tissue undergoing
CC angiogenesis. The method and the polynucleotide sequences of the
CC invention are useful for diagnosing and treating angiogenesis and
CC angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC sequences are also useful in the gene therapy of such disorders. The
CC angiogenesis-associated proteins encoded by the polynucleotide sequences
CC are useful as a vaccine for therapeutic and prophylactic immunisation.
CC ABU03456-ABU03569 represent angiogenesis-associated protein sequences
XX
SQ Sequence 1170 AA;

Query Match          67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELANL 13
       : ||||| | :
Db      301 VTEENKELANLNR 313

RESULT 46
ABG74673
ID ABG74673 standard; protein; 1170 AA.
XX
AC ABG74673;
XX
XX 10-MAY-2003 (first entry)
XX
DE Human THBS1 protein.
XX
KW Human; THBS1; vascular disease; cardiant; antiarteriosclerotic; stroke;
KW cerebroprotective; gene therapy; coronary artery disease; ischaemia;
KW myocardial infarction; peripheral vascular disease; pulmonary embolism;
KW venous thromboembolism; forensic; paternity testing.
XX
OS Homo sapiens.
XX
PN WO2003016494-A2.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US026343.
XX
PR 16-AUG-2001; 2001US-0313097P.
XX
PR 05-OCT-2001; 2001US-0327485P.
XX
PR 14-DEC-2001; 2001US-00020141.
XX
PA (VITI-) VITIVITY INC.
XX
PI McCarthy J, Ableson A;
XX
DR WPI; 2003-300617/29.
XX
DR N-PSDB; ABQ77405.
XX
PT Identifying a subject as a candidate for a particular course of therapy
PT to treat a vascular disease or disorder, e.g. stroke, myocardial
PT infarction or ischemia by determining the identity of the nucleotide
PT present at specific positions.
XX
PS Disclosure; Fig 12; 568pp; English.
XX
CC This invention describes a novel method for identifying a subject as a
CC candidate for a particular course of therapy to treat a vascular disease
CC or disorder. The method comprises determining the identity of the
CC nucleotide present at specific positions, or their complements, and
CC identifying the subject as a candidate for a particular clinical course
CC of therapy based on the identity of the nucleotide present in that
CC specific position. The method can be used for identifying a subject who
CC is a candidate for further diagnostic evaluation of a vascular disease or
CC disorder and selecting a clinical course of therapy. The products of the
CC invention have cardiant, antiarteriosclerotic and cerebroprotective
CC activity and can be used for gene therapy. The methods disclosed are
CC useful for treating a vascular disease, e.g. atherosclerosis, coronary
CC artery disease, myocardial infarction, ischaemia, stroke, peripheral
CC vascular diseases, venous thromboembolism and pulmonary embolism. The DNA
CC sequences are useful as fingerprint for detecting different individuals
CC within the same species applicable in forensic studies and paternity
CC testing. This sequence represents the human THBS1 gene represented in
CC GI12583762, used to illustrate the method of the invention
XX
SQ Sequence 1170 AA;

Query Match          67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELANL 13
       : ||||| | :
Db      301 VTEENKELANLNR 313

RESULT 47
AAE36228
```

ID AAE36228 standard; protein; 1170 AA.
XX
AC AAE36228;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human THBS1 reference protein.
XX
KW Human; thrombospondin 1; THBS1; vascular disease; therapy; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 700
FT /note= "This residue is changed to Ser during single
FT nucleotide polymorphism (SNP)"
XX
PN WO2003020120-A2.
XX
PD 13-MAR-2003.
XX
PF 04-SEP-2002; 2002WO-US028281.
XX
PR 04-SEP-2001; 2001US-0317033P.
PR 17-OCT-2001; 2001US-0330248P.
PR 14-DEC-2001; 2001US-00017721.
XX
XX (VITI-) VITIVITY INC.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI McCarthy J, Daley G, Bolk S;
XX
XX WPI; 2003-312992/30.
XX
PT Novel isolated nucleic acid comprising allelic variant of polymorphic
PT region of thrombospondin 1 or 4 gene, useful for identifying
PT predisposition of subject to develop vascular disease associated with the
PT allelic variant.
XX
PS Example 1; Fig 2; 338pp; English.
XX
XX The invention is based at least in part on the discovery of polymorphisms
CC within the thrombospondin 1 (THBS1) or THBS4 genes. The invention also
CC provides a nucleic acid comprising allelic variant of polymorphic region
CC of THBS1 or THBS4 gene. The method is useful for identifying a subject as
CC a candidate for a particular clinical course of therapy or to treat a
CC vascular disease or disorder and for selecting a clinical course of
CC therapy to treat a subject who is at risk for developing a vascular
CC disease or disorder. The present sequence is human THBS1 reference
CC protein. Note: This sequence is said to be encoded by SEQ ID NO: 1
CC (AAD55693). However this does not appear to be the case
XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB 301 VTEENKELANEELR 313
RESULT 48
ABR62059
ID ABR62059 standard; protein; 1170 AA.
XX
AC ABR62059;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human thrombospondin 1 (THBS1).

XX TSPI; immunogenic; CD36; immunomodulator; vaccine; gene therapy;
KW immune response; human; thrombospondin 1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT /note= "signal sequence"
FT Protein 32..1170
FT /note= "mature protein"
FT Domain 447..452
FT /note= "CD36 binding domain"
FT Domain 504..511
FT /note= "CD36 binding domain"
XX
PN WO2003050268-A2.
XX
PD 19-JUN-2003.
XX
PF 12-DEC-2002; 2002WO-US039885.
XX
PR 12-DEC-2001; 2001US-0341771P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Cox WI, Alexander JP, Goebel S;
XX
XX WPI; 2003-513975/48.
DR N-PSDB; ACC84978.
XX
XX New nucleic acid molecule comprising first nucleic acid sequence encoding
PT at least one CD36 binding domain, and a second nucleic acid sequence
PT encoding at least one immunogenic amino acid sequence, useful for
PT eliciting immune response.
XX
PS Example 1; Fig 1A; 59pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding an
CC immunogenic chimeric polypeptide. The isolated nucleic acid molecule
CC comprises at least a first nucleic acid sequence encoding at least one
CC CD36 binding domain, and a second nucleic acid sequence encoding at least
CC one immunogenic amino acid sequence, where the introduction of the
CC nucleic acid molecule in a cell results in expression of the immunogenic
CC chimeric polypeptide. The nucleic acid molecule is useful for eliciting
CC an immune response. The present sequence represents a human
CC thrombospondin 1 (TSPI) polypeptide
XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 7; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB 301 VTEENKELANEELR 313
RESULT 49
ADN39852
ID ADN39852 standard; protein; 1170 AA.
XX
AC ADN39852;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO: C222.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX Homo sapiens.
XX WO20003042661-A2.
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 09-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372248P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39634.
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO C222; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX Sequence 1170 AA;
SQ
Query Match 67.7%; Score 42; DB 7; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILLENKELENAALK 13
Db 301 VTEENKELANEUR 313

RESULT 50
ADJ75296
ID ADJ75296 standard; protein; 1170 AA.
XX AC ADJ75296;
XX 20-MAY-2004 (first entry)
XX DE
XX Marker gene related amino acid sequence SEQ ID NO:548.
XX bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker.
XX Homo sapiens.
XX EPI394274-A2.
XX 03-MAR-2004.
XX 04-AUG-2003; 2003EP-00254857.
XX 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX (GENO-) GENOX RES INC.
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
XX WPI; 2004-193155/19.
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX Example 11; SEQ ID NO 548; 241pp; English.
XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX Sequence 1170 AA;
SQ
Query Match 67.7%; Score 42; DB 8; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;

XX	
XX	
PT	New purified thrombospondin fragment extracted from a body fluid, useful
PT	for diagnosing cancer e.g. adenoma, adenocarcinoma, carcinoma, lymphoma
PT	or leukemia or as calibrators, indicators, immunogens and analytes.
XX	
XX	
PS	Disclosure; SEQ ID NO 38; 76pp; English.
XX	
XX	
CC	The present sequence is that of full-length human thrombospondin-1 (TSP)
CC	The invention relates to TSP fragments (80-100, 40-55 or 20-35 kDa

CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX
SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 8; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 56
ADQ39356
ID ADQ39356 standard; protein; 1170 AA.

XX AC ADQ39356;
XX DT 18-NOV-2004 (first entry)
XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1019.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human.
XX OS Homo sapiens.

XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.

XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.

XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38528.

XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.

XX PS Claim 10; SEQ ID NO 1019; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The

CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 1170 AA;
Query Match 67.7%; Score 42; DB 8; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 57
ADQ39355
ID ADQ39355 standard; protein; 1170 AA.

XX AC ADQ39355;
XX DT 18-NOV-2004 (first entry)
XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1018.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human.
XX OS Homo sapiens.

XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.

XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.

XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX DR WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38527.

XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.

XX PS Claim 10; SEQ ID NO 1018; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the

CC	CC	CC	CC	CC	CC	XX	XX	SQ	Query Match	67.7%	Score 42;	DB 9;	Length 1170;	
									Best Local Similarity	61.5%	Pred. No. 9.8e+02;			
									Matches	8;	Conservative	2;	Mismatches	3; Indels
														0; Gaps
														0;
Qy									1	ILEENKELENALK	13			
									:		:			
Db									301	VTEENKELANELR	313			

RESULT	59	
ABB87781		
ID	ABB87781	standard; protein; 1170 AA.
XX	XX	
XX	ABB87781;	
XX	XX	
XX	XX	
DT	06-OCT-2005	(first entry)
XX	XX	
DE	Human thrombospondin 1,	breast tumor marker.
XX	XX	
XX	XX	
KW	Breast tumor; breast disease;	endocrine disease;

KW	DNA library; microarray; expression; thrombospondin 1.
XX	
XX	Homo sapiens.
OS	
XX	
XX	WO2005071419-A2.
XX	
XX	
XX	04-AUG-2005.
PD	
XX	
XX	17-JAN-2005; 2005WO-IB000261.
XX	
XX	
PR	16-JAN-2004; 2004US-0537412P.
XX	
PR	14-JAN-2005; 2005US-00036298.
XX	
PA	(IPSO-) IPSOGEN.
PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA	(PAOL-) INST PAOLI CALMETTES.
XX	
PI	Jacquemier J, Bertucci F, Birnbaum D, Debono S, Tagett R;
XX	
DR	WPI; 2005-564257/57.

XX Analyzing differential protein expression associated with histopathologic
PT features of breast disease, comprises detecting overexpression or
PT underexpression of pool of proteins having Afadin, Aurora A, in breast
PT tissues or cells.
XX
XX Disclosure; SEQ ID NO 51; 87pp; English.
PS
XX The present invention provides a method for analyzing differential
CC protein expression associated with histopathological features of breast
CC disease, in particular breast tumors, e.g. breast carcinomas, comprising
CC

breast tissues or cells. The pool comprises all or part of a protein set comprising: afadin, aurora A, alpha-catenin, beta-catenin, BCL2, cyclin D1, cyclin E, cytokeratin 5/6, cytokeratin 8/18, E-cadherin, EGFR, ERBB2, ERBB3, ERBB4, estrogen receptor, FGFR1, FHIT, GATA3, K167, mucin 1, p53, P-cadherin, progesterone receptor, TACC1, TACC2, TACC3, cytokeratin 6, cytokeratin 18, angrl, aurora B, BCRP1, cathepsin D, CD10, CD44, CK14, CK2, fibroblast growth factor 2 (FGF2), GATA4, hfla, matrix metalloproteinase (MMP) 9, MTAL, NM23, neueregulins (NRG) 1 alpha, NRG1 beta, p27, parkin, PLAU, S100, SCRIBBLE, smooth muscle actin, thrombospondin (THBS) 1, tissue-inhibitors of matrix metalloproteinase 1 (TIMP1). The measuring of over- or underexpression of proteins is carried out on tissue microarray, by immunohistochemistry (IHC) technologies. The method may involve comparing expression levels of the protein set in a

CC control sample to levels of equivalent proteins in a tissue sample. The
 CC detection may also involve detecting the over- or underexpression of
 CC nucleic acids coding for the proteins. Also claimed are a protein library
 CC and a nucleic acid library useful for the molecular characterization of
 CC histopathological features of breast disease. The method is useful for
 CC analyzing differential protein expression associated with
 CC histopathological features of breast disease. It is also useful for
 CC detecting, diagnosing, staging, monitoring, predicting, and preventing
 CC conditions associated with breast cancer, predicting clinical outcome of
 CC breast cancer, predicting occurrence of metastatic relapse, and for
 CC determining the stage or aggressiveness of a breast cancer. Treatment of
 CC a patient can be based on the analysis of the differential protein
 CC expression profile. Breast cancers can be classified into prognostically
 CC relevant subclasses, and appropriate doses and/or schedule of
 CC chemotherapeutics and/or biopharmaceuticals and/or radiation therapy can
 CC be selected to circumvent toxicities in a patient. The method is also
 CC useful in assessing breast cancer heterogeneity and prognosis in patients
 CC with stage I, II or III disease. The present sequence is that of
 CC thrombospondin 1 (THBS1).
 XX
 SQ Sequence 1170 AA;

Query Match 67.7%; Score 42; DB 9; Length 1170;
 Best Local Similarity 61.5%; Pred. No. 9.8e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 : ||||| | :
 DB 301 VTEENKELANELR 313

RESULT 60

AEBA6751

ID AEB46751 standard; protein; 1170 AA.

XX AC AEB46751;

XX DT 20-OCT-2005 (first entry)

XX DE Human thrombospondin 1.

XX KW Thrombospondin; gastrointestinal proliferative factor; GIPF; mucositis;
 KW enterocolitis; inflammatory bowel disease; Crohns disease;
 KW short-bowel syndrome; proliferation; epithelial cell; wound healing;
 KW burns; ocular disease; ophthalmological; antiinflammatory;
 KW gastrointestinal-gen.; vulnery; immunosuppressive; chemotherapy;
 KW radiotherapy; B-lymphocyte; transgenic animal.

XX OS Homo sapiens.

XX PN WO2005072419-A2.

XX PD 11-AUG-2005.

XX PF 27-JAN-2005; 2005WO-US002996.

XX PR 27-JAN-2004; 2004US-0539605P.

XX PR 15-OCT-2004; 2004US-0619241P.

XX (NUVE-) NUVELO INC.

XX (KIRI) KIRIN BEER KK.

XX Boyle BJ, Funk W, Kakitani M, Oshima T, Park EJ, Tang TY;

XX Yagi M, Tomizuka K;

XX WPI; 2005-555619/56.

XX New composition comprising a gastrointestinal proliferative factor (GIPF)
 PT polypeptide, its fragment or analogue, and a carrier, useful for treating
 PT mucositis, inflammatory bowel disease or short bowel syndrome.

PS Disclosure; SEQ ID NO 28; 377pp; English.

XX

CC The invention relates to a composition comprising a gastrointestinal
 CC proliferative factor (GIPF) polypeptide, its fragment or analogue, and a
 CC carrier. Also included are a pharmaceutical composition (comprising a
 CC polypeptide comprising a biologically active fragment of GIPF and a
 CC carrier), a method of stimulating epithelial cell proliferation in a
 CC subject (comprising administering to the subject the composition), a
 CC method of treating a disease (chosen from mucositis, inflammatory bowel
 CC disease or short bowel syndrome, comprising administering to the
 CC mammalian subject the composition), a method for treating a patient at
 CC risk for damage to epithelial cells lining at least a portion of the
 CC gastrointestinal tract by administering the composition, a method for
 CC treating a patient that has undergone radiation therapy/chemotherapy by
 CC administering the composition, an adenoviral vector (comprising a gene
 CC encoding GIPF operably associated with an expression control sequence), a
 CC pharmaceutical composition comprising the adenoviral vector, a transgene
 CC construct comprising a nucleic acid encoding a GIPF protein (where the
 CC nucleic acid is operably linked to transcriptional regulatory sequences
 CC directing its expression in B-cells), a transgenic mouse that produces in
 CC its B-cells detectable levels of a native human GIPF protein (where the
 CC transgenic mouse has stably integrated into its genome a nucleic acid
 CC sequence encoding GIPF protein, operably linked to transcriptional
 CC regulatory sequences directing its expression to B-cells), a method of
 CC identifying a drug candidate (for treating mucositis, inflammatory bowel
 CC disease, or short bowel syndrome), an isolated polynucleotide encoding
 CC the polypeptide, an expression vector comprising expression regulatory
 CC elements operatively linked to the polynucleotide, a host cell
 CC transformed/transfected with the polynucleotide, methods of producing the
 CC polypeptide/pharmaceutical composition and a polypeptide (comprising a
 CC variant of SEQ ID No. 4, where the valine amino acid 50 is replaced with
 CC isoleucine). The composition is useful for treating mucositis,
 CC inflammatory bowel disease, short bowel syndrome, wounds, burns, ocular
 CC diseases and for treating a patient at risk for damage to epithelial
 CC cells lining at least a portion of the gastrointestinal tract, or a
 CC patient that has undergone radiation therapy or chemotherapy by
 CC administering the composition. The present sequence represents a human
 CC protein showing some sequence similarity to GIPF.

XX SQ Sequence 1170 AA;

Query Match 67.7%; Score 42; DB 9; Length 1170;

Best Local Similarity 61.5%; Pred. No. 9.8e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 301 VTEENKELANELR 313

Search completed: March 11, 2006, 12:18:52

Job time : 138.447 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:19:28 ; Search time 20.5263 Seconds
(without alignments)
60.937 Million cell updates/sec

Title: US-10-774-242A-6

Perfect score: 62

Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : PIR 80:**

1: Pir1:
2: Pir2:
3: Pir3:
4: Pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	759	2 I38593	fibroblast activat
2	47	75.8	405	2 T41194	probable high-affi
3	45	72.6	162	2 H90365	hypothetical prote
4	43	69.4	706	2 S62933	hypothetical prote
5	43	69.4	806	2 A71979	hypothetical prote
6	43	69.4	831	2 B64528	conserved hypothet
7	42	67.7	172	2 C71038	hypothetical prote
8	42	67.7	877	2 H71647	alanine-tRNA ligas
9	42	67.7	1170	1 TSHUP1	thrombospondin 1 p
10	41	66.1	291	2 T01241	probable MYB fami
11	41	66.1	294	2 AF1218	Salmonella typhimu
12	41	66.1	546	2 G72210	hypothetical prote
13	41	66.1	847	2 A64675	alanine-tRNA ligas
14	41	66.1	847	2 F71842	alanine-tRNA ligas
15	40	64.5	245	2 S71116	BCR-associated prot
16	40	64.5	340	2 A53378	recombination prot
17	40	64.5	380	2 H72374	probable phosphori
18	40	64.5	381	2 G95085	conserved hypothet
19	40	64.5	397	2 C97953	conserved hypothet
20	40	64.5	959	2 A55913	transcytosis-assoc
21	40	64.5	1173	2 T43527	sp8 protein - firs
22	40	64.5	1727	2 T50073	myosin-like coiled
23	40	64.5	2285	2 T12796	probable transglyc
24	40	64.5	2561	2 T24864	hypothetical prote
25	39	62.9	168	2 H64034	hypothetical prote
26	39	62.9	171	2 E97074	probable acetyltra
27	39	62.9	230	2 A86297	hypothetical prote
28	39	62.9	291	2 G86903	cell shape determi
29	39	62.9	405	2 D64522	carboxynorpermid

30	39	62.9	405	2 H71984	probable carboxyno
31	39	62.9	442	2 B71083	hypothetical prote
32	39	62.9	453	2 F75206	maltose-binding pe
33	39	62.9	609	2 T02769	Gns protein - Adel
34	39	62.9	629	2 A71023	arginine-tRNA liga
35	39	62.9	656	2 AH2402	cell division prot
36	39	62.9	812	2 A81621	conserved hypothet
37	39	62.9	831	2 F72044	ct620 hypothetical
38	39	62.9	831	2 C86581	CT620 hypothetical
39	39	62.9	878	2 G97865	alanine-tRNA ligas
40	39	62.9	991	2 A99604	hypothetical prote
41	39	62.9	3225	2 I52300	giantin - human
42	39	62.9	3259	1 A56539	giantin - human
43	38.5	62.1	453	2 H70333	conserved hypothet
44	38	61.3	82	2 A60457	clusterin - sheep
45	38	61.3	133	2 JN0729	hypothetical 16.2K
46	38	61.3	276	2 A64516	hypothetical prote
47	38	61.3	294	2 A11571	Regulatory protein
48	38	61.3	310	2 A80078	LY8R-family regula
49	38	61.3	350	2 T33458	hypothetical prote
50	38	61.3	373	2 G71917	8-amino-7-oxonon
51	38	61.3	395	2 D72339	hypothetical prote
52	38	61.3	430	2 D72324	O-acetylhomoserine
53	38	61.3	439	2 A35744	clusterin precurs
54	38	61.3	445	2 A40018	clusterin precurs
55	38	61.3	446	2 A42108	clusterin precurs
56	38	61.3	448	2 C81718	signal recognition
57	38	61.3	448	2 D71566	probable signal re
58	38	61.3	527	2 T28878	hypothetical prote
59	38	61.3	575	2 S22594	hypothetical prote
60	38	61.3	575	2 S22595	hypothetical prote
61	38	61.3	609	2 S67616	probable two-compo
62	38	61.3	612	2 F83000	hypothetical prote
63	38	61.3	669	2 A72718	hypothetical prote
64	38	61.3	762	2 E64597	DNA mismatch repai
65	38	61.3	762	2 H71916	DNA mismatch repai
66	38	61.3	843	1 A27340	complement C7 prec
67	38	61.3	910	2 E89918	2-oxoglutarate de
68	37.5	60.5	279	2 C81412	NOL1/NOP2/sun fami
69	37	59.7	77	2 B47074	orf 3' of rfap - S
70	37	59.7	102	2 T31879	hypothetical prote
71	37	59.7	102	2 T45810	hypothetical prote
72	37	59.7	130	2 AH1125	hypothetical prote
73	37	59.7	137	1 S38484	urease (BC 3.5.1.5
74	37	59.7	168	2 D82833	transcription regu
75	37	59.7	227	2 S74350	hypothetical prote
76	37	59.7	255	1 S14051	ribosomal protein
77	37	59.7	255	2 S21121	ribosomal protein
78	37	59.7	294	2 G85474	homeodomain-like p
79	37	59.7	301	2 T06112	homeotic protein r
80	37	59.7	310	2 E84612	homeodomain transc
81	37	59.7	311	2 S47136	homeotic protein A
82	37	59.7	342	2 S50400	hypothetical prote
83	37	59.7	348	2 T51371	hypothetical prote
84	37	59.7	364	2 T03892	hypothetical prote
85	37	59.7	371	2 D97042	hypothetical prote
86	37	59.7	397	2 D64379	dihydrolipoamide d
87	37	59.7	451	2 A53910	spindle pole body
88	37	59.7	466	2 A13262	argininosuccinate
89	37	59.7	502	2 T52157	hypothetical prote
90	37	59.7	502	2 A83938	lipopolysaccharide
91	37	59.7	522	2 C96608	hypothetical prote
92	37	59.7	588	2 P00007	hypothetical prote
93	37	59.7	862	2 T01798	type IIS restricti
94	37	59.7	1279	2 E64709	replication factor
95	37	59.7	1847	2 E64477	hypothetical prote
96	37	59.7	2216	2 S78398	kinesin-related pr
97	37	59.7	2954	2 T14156	hypothetical prote
98	37	59.7	3394	2 T18501	hypothetical prote
99	36	58.1	82	2 D64329	hypothetical prote
100	36	58.1	101	2 B96904	hypothetical prote
101	36	58.1	126	2 S49392	hypothetical prote
102	36	58.1	152	2 S03630	neurogenic gene co

103	36	58.1	155	2	C97052	chorismate mutase	176	35	56.5	195	2	D71872	hypothetical prote
104	36	58.1	180	2	T29403	hypothetical prote	177	35	56.5	195	2	A48957	orf 2 5' of pepC -
105	36	58.1	191	2	G82907	conserved hypothet	178	35	56.5	197	2	A48957	hypothetical prote
106	36	58.1	201	2	B71861	hypothetical prote	179	35	56.5	198	2	E86183	hypothetical prote
107	36	58.1	202	2	G64565	hypothetical prote	180	35	56.5	204	2	S51232	gibberellin-respon
108	36	58.1	213	2	T22984	hypothetical prote	181	35	56.5	218	2	F64461	hypothetical prote
109	36	58.1	245	2	C34787	tropomyosin 3 alph	182	35	56.5	218	2	D97536	hypothetical prote
110	36	58.1	261	2	B55590	hypothetical prote	183	35	56.5	230	2	140287	outer surface prot
111	36	58.1	265	2	C70221	conserved hypothet	184	35	56.5	234	2	T06995	probable MADS box
112	36	58.1	268	2	C84957	elongation factor	185	35	56.5	240	2	S61910	hypothetical prote
113	36	58.1	273	2	H64349	hypothetical prote	186	35	56.5	247	2	T48878	proteasome pma, a
114	36	58.1	281	2	A34787	tropomyosin 1 alph	187	35	56.5	262	2	G75039	hypothetical prote
115	36	58.1	297	2	T28785	hypothetical prote	188	35	56.5	265	2	T28102	hypothetical prote
116	36	58.1	326	2	A70368	conserved hypothet	189	35	56.5	290	2	T16212	conserved hypothet
117	36	58.1	327	2	T23572	hypothetical prote	190	35	56.5	305	2	D70432	hypothetical prote
118	36	58.1	347	2	D69166	N5,N10-methylenete	191	35	56.5	312	2	A64461	hypothetical prote
119	36	58.1	357	2	T15743	hypothetical prote	192	35	56.5	326	2	T29537	hypothetical prote
120	36	58.1	392	2	T04150	RAB23 protein homo	193	35	56.5	332	2	C64183	hypothetical prote
121	36	58.1	420	2	S06652	hypothetical prote	194	35	56.5	336	2	T44795	methyltransferase
122	36	58.1	435	2	G46666	conserved hypothet	195	35	56.5	336	2	H71103	hypothetical prote
123	36	58.1	435	2	B71850	hypothetical prote	196	35	56.5	347	2	H90586	oxygenase (importe
124	36	58.1	444	2	T42537	probable transketo	197	35	56.5	357	2	I55210	crucicboxylase car
125	36	58.1	454	2	S43556	plasmaonogen-bindin	198	35	56.5	358	2	S73776	MG269 homolog Fil
126	36	58.1	491	2	S43678	killer toxin - Pic	199	35	56.5	361	1	F64355	conserved hypothet
127	36	58.1	497	2	A84641	hypothetical prote	200	35	56.5	371	2	C95130	conserved hypothet
128	36	58.1	509	2	A40448	DNA-binding protei	201	35	56.5	371	2	F81102	hypothetical prote
129	36	58.1	532	2	S54871	M protein - Strept	202	35	56.5	371	2	D81838	hypothetical prote
130	36	58.1	553	2	F81282	probable ferredoxi	203	35	56.5	371	2	T05376	hypothetical prote
131	36	58.1	582	2	D96903	hydrogene dehydrog	204	35	56.5	375	2	B98001	glycerate kinase [
132	36	58.1	582	2	JC5002	hydrogenase [EC 1.	205	35	56.5	376	2	G71290	flagellar biosynth
133	36	58.1	587	2	T01829	hypothetical prote	206	35	56.5	376	2	G71828	hypothetical prote
134	36	58.1	599	2	T07733	probable chaperoni	207	35	56.5	382	2	E64686	myosin-like protei
135	36	58.1	602	1	TVRTRR	protein kinase (EC	208	35	56.5	433	2	T43924	anchraniolate synth
136	36	58.1	613	2	T43933	DNA-directed DNA p	209	35	56.5	435	2	C70456	homoserine dehydro
137	36	58.1	621	2	A37103	lamin precursor -	210	35	56.5	435	2	G64305	hypothetical prote
138	36	58.1	646	2	G84854	hypothetical prote	211	35	56.5	464	2	F69587	L-arabinose transp
139	36	58.1	685	2	T40162	transketolase - fi	212	35	56.5	483	2	I40716	H+-transporting tw
140	36	58.1	714	2	G71963	fumarate reductase	213	35	56.5	490	2	E69066	conserved hypothet
141	36	58.1	737	2	D71122	hypothetical prote	214	35	56.5	502	2	T45852	hypothetical prote
142	36	58.1	761	2	D70447	tetrahydropteroyl	215	35	56.5	508	2	E71620	hypothetical prote
143	36	58.1	777	2	S58162	probable Rho GTPas	216	35	56.5	526	1	P5XRBT	nickel ABC transpo
144	36	58.1	780	2	G84708	probable VPI/ABI3	217	35	56.5	528	2	D83874	sensor histidine k
145	36	58.1	802	2	T21315	hypothetical prote	218	35	56.5	532	2	E75617	hypothetical prote
146	36	58.1	838	2	JC7363	95K retinoblastoma	219	35	56.5	574	2	S22596	DNA-directed RNA p
147	36	58.1	918	2	E90542	lipoprotein (impor	220	35	56.5	604	2	S02195	hypothetical prote
148	36	58.1	933	2	T28995	hypothetical prote	221	35	56.5	612	2	E84809	hypothetical prote
149	36	58.1	978	2	A70387	conserved hypothet	222	35	56.5	615	2	T06108	hypothetical prote
150	36	58.1	1039	2	T28644	Y4JQ protein - Rhi	223	35	56.5	625	2	H75110	arginyl-tRNA synth
151	36	58.1	1170	2	A40558	thrombospondin 1 p	224	35	56.5	646	1	S15901	chromogranin B pre
152	36	58.1	1181	2	T19736	hypothetical prote	225	35	56.5	646	2	T41545	hypothetical trans
153	36	58.1	1306	2	T28313	ORF MSV152 probabl	226	35	56.5	647	2	T30892	hypothetical prote
154	36	58.1	1726	2	A39401	merozoite surface	227	35	56.5	658	2	C82882	ABC transporter UU
155	36	58.1	1738	2	C84507	hypothetical prote	228	35	56.5	665	2	C71667	propionyl-CoA carb
156	36	58.1	1751	2	A45604	major blood-stage	229	35	56.5	677	1	S09078	chromogranin B pre
157	36	58.1	1799	2	C71622	hypothetical prote	230	35	56.5	702	1	J00868	glucan 1,4-alpha-g
158	36	58.1	1983	2	T00385	KIAA0624 protei	231	35	56.5	706	2	D82160	hypothetical prote
159	36	58.1	2581	2	AF2545	hypothetical prote	232	35	56.5	714	2	H64543	funarate reductase
160	36	58.1	4388	2	T28667	dynein beta heavy	233	35	56.5	789	2	S62172	SHB4 protein - yea
161	35.5	57.1	145	2	T29389	hypothetical prote	234	35	56.5	861	2	T00434	probable kinase h
162	35	56.5	51	2	D97830	hypothetical prote	235	35	56.5	863	2	C90482	ABC transporter, A
163	35	56.5	54	2	AH1378	hypothetical prote	236	35	56.5	880	2	F75103	conserved hypothet
164	35	56.5	54	2	A11747	hypothetical prote	237	35	56.5	917	2	S07183	hypothetical prote
165	35	56.5	92	2	JC3576	hypothetical prote	238	35	56.5	919	2	B72765	hypothetical prote
166	35	56.5	145	2	E70168	hypothetical 10K p	239	35	56.5	921	2	T01775	hypothetical prote
167	35	56.5	159	2	T32043	flagellar protein	240	35	56.5	943	2	S54493	hypothetical prote
168	35	56.5	175	2	T13223	hypothetical prote	241	35	56.5	987	2	T51360	kinesin-like heavy
169	35	56.5	175	2	JC5912	protein R175 - Lac	242	35	56.5	1006	2	T02017	kinesin-related pr
170	35	56.5	178	2	C69206	Lysa protein - Lac	243	35	56.5	1025	2	T44802	type I site-specif
171	35	56.5	182	2	A53803	hypothetical prote	244	35	56.5	1032	2	B86224	hypothetical prote
172	35	56.5	188	2	B82471	dihydrofolate redu	245	35	56.5	1054	2	T30177	cytoskeleton assem
173	35	56.5	191	2	T31903	hypothetical prote	246	35	56.5	1154	2	T15650	hypothetical prote
174	35	56.5	195	2	G86742	hypothetical prote	247	35	56.5	1254	2	S46636	hypothetical prote
175	35	56.5	195	2	AE2755	conserved hypothet	248	35	56.5	1261	2	E59430	PTPL1-associated R

249	35	56.5	1289	2	S67200	hypothetical prote	322	34	54.8	373	2	F64594	8-amino-7-oxononan
250	35	56.5	1394	2	A29637	position-specific	323	34	54.8	393	2	A70319	ATP synthase F1 ga
251	35	56.5	1434	2	C82923	DNA-directed RNA p	324	34	54.8	414	2	G69333	probable acyl-CoA
252	35	56.5	1558	2	T17603	REGA-H3 antigen pF	325	34	54.8	420	2	T44132	plasmid recombinat
253	35	56.5	1854	2	T13576	hypothetical prote	326	34	54.8	424	2	T29158	hypothetical prote
254	35	56.5	2025	2	JC5020	tetratricopeptide	327	34	54.8	427	2	H69787	DNA-methyltransfer
255	35	56.5	2104	2	T38774	myosin-3 heavy cha	328	34	54.8	431	2	A72549	probable pyruvate
256	35	56.5	2116	2	A26655	myosin heavy chain	329	34	54.8	435	2	D83505	conserved hypothet
257	35	56.5	2182	2	T28634	variant-specific s	330	34	54.8	447	2	F90248	hypothetical prote
258	35	56.5	2231	2	D71870	hypothetical prote	331	34	54.8	456	2	S04079	nitrogenase (EC 1.
259	35	56.5	4550	2	T18440	hypothetical prote	332	34	54.8	470	2	T32107	hypothetical prote
260	34.5	55.6	113	2	C90279	conserved hypothet	333	34	54.8	479	2	E72254	conserved hypothet
261	34.5	55.6	401	2	G72245	aspartokinase II -	334	34	54.8	482	2	G71603	chromatin-binding
262	34	54.8	56	1	ZKBP4	gene K protein - p	335	34	54.8	505	2	A46570	H+-transporting tw
263	34	54.8	79	2	G84082	hypothetical prote	336	34	54.8	505	2	T19770	hypothetical prote
264	34	54.8	96	2	S51930	homeotic protein C	337	34	54.8	513	2	A96842	protein-export mem
265	34	54.8	108	2	B82926	hypothetical prote	338	34	54.8	518	2	E71663	hypothetical prote
266	34	54.8	113	2	F64443	hypothetical prote	339	34	54.8	526	2	T21811	hypothetical prote
267	34	54.8	125	2	T16156	hypothetical prote	340	34	54.8	539	2	D71260	hypothetical prote
268	34	54.8	151	2	S61384	icmw protein - Leg	341	34	54.8	546	2	S56306	hypothetical prote
269	34	54.8	153	2	A64369	hypothetical prote	342	34	54.8	564	2	T25945	hypothetical prote
270	34	54.8	154	2	A64097	conserved hypothet	343	34	54.8	566	2	T45626	hypothetical prote
271	34	54.8	156	2	A75034	hypothetical prote	344	34	54.8	576	2	A44085	p element transpos
272	34	54.8	160	2	B86614	CT832 hypothetical	345	34	54.8	595	2	T06412	probable chaperoni
273	34	54.8	160	2	F72010	N utilization subs	346	34	54.8	608	2	T47363	hypothetical prote
274	34	54.8	163	2	B81530	hypothetical prote	347	34	54.8	617	2	S27389	secretogranin II -
275	34	54.8	167	2	A64401	2-amino-4-hydroxy-	348	34	54.8	617	2	B70425	hypothetical prote
276	34	54.8	174	2	G83661	hypothetical prote	349	34	54.8	619	2	A02180	secretogranin II -
277	34	54.8	176	2	T32618	hypothetical prote	350	34	54.8	621	2	A57591	id-associated prot
278	34	54.8	183	2	A64248	ribosome releasing	351	34	54.8	651	2	F86563	hypothetical prote
279	34	54.8	186	2	C86861	hypothetical prote	352	34	54.8	651	2	A72060	hypothetical prote
280	34	54.8	191	2	E86499	CT360 hypothetical	353	34	54.8	677	2	C83210	methionyl-tRNA syn
281	34	54.8	191	2	H72122	hypothetical prote	354	34	54.8	683	2	S34700	probable purine nu
282	34	54.8	203	2	H81806	hypothetical prote	355	34	54.8	736	2	B82944	ribose/galactose A
283	34	54.8	211	2	E82055	sodium-type flagel	356	34	54.8	744	2	A81719	exodeoxyribonuclea
284	34	54.8	218	2	C82894	heat shock protein	357	34	54.8	747	2	E84698	hypothetical prote
285	34	54.8	222	2	S62001	ME15 protein - yea	358	34	54.8	759	2	F84662	trans-Golgi membra
286	34	54.8	224	2	A70728	hypothetical prote	359	34	54.8	782	2	S27833	rhoptyr-associated
287	34	54.8	226	2	T38305	hypothetical prote	360	34	54.8	813	1	D70176	endopeptidase La (
288	34	54.8	240	2	S48997	B-cell receptor as	361	34	54.8	830	2	T47377	hypothetical prote
289	34	54.8	241	2	C70178	hypothetical prote	362	34	54.8	841	2	G64336	large helicase rel
290	34	54.8	241	2	B71240	hypothetical prote	363	34	54.8	921	2	E86764	conserved hypothet
291	34	54.8	244	2	G64628	hypothetical prote	364	34	54.8	947	2	H84866	hypothetical prote
292	34	54.8	245	2	S24403	alpha-tropomyosin	365	34	54.8	982	2	S10340	DNA-directed RNA p
293	34	54.8	245	2	A97073	transcription regu	366	34	54.8	982	2	S00964	hypothetical prote
294	34	54.8	250	2	H64567	hypothetical prote	367	34	54.8	1039	2	C64418	isoleucine-tRNA li
295	34	54.8	252	2	D64307	5'-methylthiodeno	368	34	54.8	1039	2	S62509	probable vesicular
296	34	54.8	260	2	H82925	hypothetical prote	369	34	54.8	1044	2	T50213	probable vesicular
297	34	54.8	262	2	E64679	tryptophan synthas	370	34	54.8	1048	2	H64459	hypothetical prote
298	34	54.8	262	2	JC4109	triacylglycerol li	371	34	54.8	1056	2	C96661	kinesin-like prote
299	34	54.8	266	2	T30913	acetyltransferase (EC	372	34	54.8	1068	2	F84614	probable kinesin h
300	34	54.8	266	2	T40318	hypothetical prote	373	34	54.8	1091	2	AP2953	DNA polymerase III
301	34	54.8	267	2	T45960	hypothetical prote	374	34	54.8	1091	2	G98329	probable DNA polym
302	34	54.8	273	2	I40100	outer surface prot	375	34	54.8	1095	2	T01916	hypothetical prote
303	34	54.8	273	2	T31533	outer surface prot	376	34	54.8	1100	2	T30967	transcription acti
304	34	54.8	274	2	I40089	outer surface prot	377	34	54.8	1114	2	JH0284	125K surface anti
305	34	54.8	280	2	D69988	plant metabolite d	378	34	54.8	1131	2	F82875	hypothetical prote
306	34	54.8	281	2	F75216	hypothetical prote	379	34	54.8	1146	2	B70376	reverse gyrase - A
307	34	54.8	285	2	D69975	anti-sigma factor	380	34	54.8	1174	2	T08196	hypothetical prote
308	34	54.8	285	2	T32565	unknown in ISBc l	381	34	54.8	1207	2	AH2154	hypothetical prote
309	34	54.8	285	2	G85605	hypothetical prote	382	34	54.8	1287	2	T42658	hypothetical prote
310	34	54.8	312	2	C82974	hypothetical prote	383	34	54.8	1313	2	G82887	hypothetical prote
311	34	54.8	335	2	S27674	quinolinate synthe	384	34	54.8	1402	2	S62557	probable calcium-t
312	34	54.8	336	2	D64689	hypothetical prote	385	34	54.8	1610	2	A46227	voltage-dependent
313	34	54.8	336	2	A71828	quinolinate synthe	386	34	54.8	1777	2	T00490	nonstructural prot
314	34	54.8	337	2	T30511	hypothetical prote	387	34	54.8	1809	2	C97038	phage-related prot
315	34	54.8	344	2	G90968	N5,N10-methylenete	388	34	54.8	2052	2	JC5837	364K Golgi complex
316	34	54.8	352	2	D90390	conserved hypothet	389	34	54.8	3187	2	JC5837	lipophorin - fruit
317	34	54.8	358	2	AH1376	hypothetical prote	390	34	54.8	3351	1	T13812	genome polyprotein
318	34	54.8	361	1	IBMSN	hydroxymethylbilan	391	34	54.8	3386	1	GNWVDF	conserved hypothet
319	34	54.8	361	1	IBRTE	gentisate 1,2-diox	392	33.5	54.0	109	2	B72213	hypothetical prote
320	34	54.8	371	2	B83900	hypothetical prote	393	33.5	54.0	558	2	A75216	molybdenum cofacto
321	34	54.8	371	2	H70326		394	33.5	54.0	620	1	F64410	

395 33.5 54.0 1236 1 VHWWE
396 33 53.2 69 2 H70131
397 33 53.2 81 2 E90605
398 33 53.2 88 2 A72222
399 33 53.2 92 2 T31029
400 33 53.2 102 2 T32603
401 33 53.2 108 2 I48930
402 33 53.2 116 2 A89901
403 33 53.2 122 2 G97899
404 33 53.2 127 2 G84002
405 33 53.2 129 2 T17322
406 33 53.2 131 2 T25705
407 33 53.2 133 2 F68616
408 33 53.2 136 2 G71601
409 33 53.2 138 2 S76339
410 33 53.2 149 2 F69143
411 33 53.2 157 2 B64245
412 33 53.2 158 2 T15036
413 33 53.2 159 2 T10826
414 33 53.2 159 2 T27831
415 33 53.2 160 2 A90533
416 33 53.2 165 2 C95036
417 33 53.2 166 2 T39620
418 33 53.2 170 2 T48362
419 33 53.2 173 2 D86825
420 33 53.2 173 2 H97231
421 33 53.2 176 2 S06638
422 33 53.2 183 1 D64430
423 33 53.2 191 2 C64001
424 33 53.2 193 2 H84958
425 33 53.2 197 2 G95020
426 33 53.2 197 2 E97892
427 33 53.2 197 2 C69778
428 33 53.2 200 2 H82570
429 33 53.2 209 2 C90188
430 33 53.2 211 2 T00611
431 33 53.2 215 2 T23195
432 33 53.2 215 2 A11164
433 33 53.2 215 2 I15223
434 33 53.2 215 2 I52523
435 33 53.2 221 2 S72318
436 33 53.2 224 2 S85900
437 33 53.2 228 2 B86276
438 33 53.2 237 2 A45929
439 33 53.2 239 1 R3R23
440 33 53.2 239 2 D39283
441 33 53.2 241 1 D71233
442 33 53.2 243 2 E71966
443 33 53.2 250 2 D72333
444 33 53.2 251 2 A9183
445 33 53.2 253 2 B64541
446 33 53.2 253 2 C81393
447 33 53.2 254 2 A64437
448 33 53.2 262 2 H72107
449 33 53.2 262 2 B86516
450 33 53.2 264 2 C82944
451 33 53.2 271 2 A10532
452 33 53.2 271 2 G85504
453 33 53.2 271 2 G90653
454 33 53.2 271 2 E64744
455 33 53.2 271 2 AG0131
456 33 53.2 271 2 C83715
457 33 53.2 273 2 A64238
458 33 53.2 273 2 F97244
459 33 53.2 275 2 D72801
460 33 53.2 279 2 D71453
461 33 53.2 280 2 C82490
462 33 53.2 293 2 A96946
463 33 53.2 297 2 C96524
464 33 53.2 299 1 F70330
465 33 53.2 300 2 T32697
466 33 53.2 301 1 S65681
467 33 53.2 306 2 E81230

structural polypro
ribosomal protein
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
cofactor A - mouse
ribosome-binding f
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
ribosomal protein
pilin repressor pi
heat shock protein
probable resistanc
hypothetical prote
hypothetical prote
PTS system, IIA co
hypothetical prote
pyrimidine operon
probable membrane
calpastatin - bovi
probable transcrip
hypothetical prote
phosphopeptase iso
Holliday junction
holliday junction
hypothetical prote
heat shock protein
3-octaprenyl-4-hyd
hypothetical prote
hypothetical prote
probable fibronect
nucleoporin p62 ho
ribosomal protein
hypothetical prote
hypothetical prote
hypothetical prote
thiamin biosynthes
tRNA pseudouridine
hypothetical prote
probable transcrip
probable 3-isoprop
hypothetical prote
hypothetical prote
peptide ABC transp
oligopeptide trans
DNA polymerase III
probable lipoprote
probable lipoprote
probable lipoprote
probable lipoprote
probable lipoprote
positive regulator
lipase-esterase li
TPR-repeat-contain
gpl5 protein - Myc
hypothetical prote
probable potassiu
N-dimethylarginine
Fla17.4 imported
probable arsenical
hypothetical prote
peptidylprolyl iso
transcription regu

468 33 53.2 306 2 H82001
469 33 53.2 306 2 A69670
470 33 53.2 306 2 C70410
471 33 53.2 320 2 H84092
472 33 53.2 320 2 T32125
473 33 53.2 324 2 A97036
474 33 53.2 330 1 YWBS
475 33 53.2 337 2 JN0559
476 33 53.2 338 2 E95028
477 33 53.2 340 2 JX0292
478 33 53.2 340 2 C55020
479 33 53.2 340 2 T33313
480 33 53.2 341 2 T13094
481 33 53.2 347 2 AE1254
482 33 53.2 347 2 B69952
483 33 53.2 348 2 F83865
484 33 53.2 349 2 G36470
485 33 53.2 353 2 JCT661
486 33 53.2 353 2 S71965
487 33 53.2 353 2 T50482
488 33 53.2 354 2 A86843
489 33 53.2 358 2 S40709
490 33 53.2 359 2 S67691
491 33 53.2 364 2 T22807
492 33 53.2 369 2 G71902
493 33 53.2 370 2 T32615
494 33 53.2 371 2 T01303
495 33 53.2 373 2 S47455
496 33 53.2 388 2 F70133
497 33 53.2 391 2 AB0443
498 33 53.2 392 2 AB0443
499 33 53.2 395 2 D70475
500 33 53.2 395 2 G71811
500 33 53.2 395 2 B64709

probable hydrogen
choline ABC transp
hypothetical prote
phage-related prot
hypothetical prote
probable membrane
tryptophan-CRNA li
repA protein - Cam
iron ABC transport
recombination prot
hypothetical prote
major capsid prote
aspartate-semialde
conserved hypothet
hypothetical prote
Wnt-7a protein - m
G protein alpha su
GTP-binding regula
G protein alpha ch
prephenate dehydro
actin-binding prot
probable membrane
hypothetical prote
probable transamin
hypothetical prote
hypothetical prote
mitosis protein JN
flagellar-associat
hypothetical prote
conserved hypothet
transcription term
transcription term

ALIGNMENTS

RESULT 1

I38593
fibroblast activation protein-alpha - human
N:Alternate names: FAP-alpha
C:Species: Homo sapiens (man)
C>Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C:Accession: I38593
R:Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncaei, M.P.; Healey, J.; O
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A:Title: Molecular cloning of fibroblast activation protein alpha, a member of the serin
A:Reference number: I38593; MUID:94261645; PMID:7911242
A:Accession: I38593
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-759 <RES>
A:Cross-references: UNIPARC:UPI0000175B7D; EMBL:U09278; NID:G507749; PID:G507750
C:Genetics:
A:Gene: GDB:FAP
A:Cross-references: GDB:374184; OMIM:600403
A:Map position: 2q23-2q23
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: transmembrane protein

Query Match 100.0%; Score 62; DB 2; Length 759;

Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ILEENKELENALK 13

Db 487 ILEENKELENALK 499

RESULT 2

T41194
probable high-affinity nickel transport protein - fission yeast (Schizosaccharomyces pom

C;Species: Schizosaccharomyces pombe
C;Date: 20-Oct-2000 #sequence_revision 08-Dec-2000 #text_change 09-Jul-2004
C;Accession: T41194; T41592
R;Bothe, G.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21977
A;Accession: T41194
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-248 <BOT>
A;Cross-references: UNIPROT:O74869; UNIPARC:UPI00000161F0A; EMBL:AL031787; NID:G3687487;
A;Experimental source: strain 972h-; cosmid c1884
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z22003
A;Accession: T41592
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 83-405 <LYN>
A;Cross-references: UNIPARC:UPI00001690B6; EMBL:AL031825; NID:G3702633; PIDN:CAA21226.1;
A;Experimental source: strain 972h-; cosmid c757
C;Genetics:
A;Gene: SPDB:SPCC1884.02; SPDB:SPCC757.01
A;Map position: 3
C;Superfamily: nickel transport protein YPO2673

Query Match 75.8%; Score 47; DB 2; Length 405;
Best Local Similarity 69.2%; Pred. No. 5.1;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||:|||||
Db 371 VLENNKELENVSK 383

RESULT 3
H90365
hypothetical protein SS01993 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: H90365
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: H90365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: UNIPROT:Q97WW6; UNIPARC:UPI0000064631; GB:AE006641; NID:G13815271; F
C;Genetics:
A;Gene: SS01993

Query Match 72.6%; Score 45; DB 2; Length 162;
Best Local Similarity 66.7%; Pred. No. 4.1;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|||||:|:
Db 24 LEENKKIENEIK 35

RESULT 4
S62933
hypothetical protein YNL021w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein N2819
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C;Accession: S62933
R;Andre, B.; Iraqi Housaini, I.; Urrestarazu, L.A.; Visseers, S.
submitted to the Protein Sequence Database, April 1996

A;Reference number: S62920
A;Accession: S62933
A;Molecule type: DNA
A;Residues: 1-706 <AND>
A;Cross-references: UNIPROT:P53973; UNIPARC:UPI00000531CD; EMBL:Z71297; NID:G1301850; PI
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:HDAL
A;Cross-references: SGD:S0004966; MIPS:YNL021w
A;Map position: 14L
F;51-388/Domain: RPD3/acuC homology <RAH1>

Query Match 69.4%; Score 43; DB 2; Length 706;
Best Local Similarity 81.8%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENAL 12
|||||:|:
Db 22 LEENKEEENSL 32

RESULT 5
A71979
hypothetical protein jhp0061 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: A71979
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho-
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: A71979
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-806 <ARN>
A;Cross-references: UNIPROT:Q9ZM29; UNIPARC:UPI00000D7152; GB:AE001445; GB:AE001439; NI
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0061

Query Match 69.4%; Score 43; DB 2; Length 806;
Best Local Similarity 69.2%; Pred. No. 48;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||:|||||
Db 155 ILTESKEIEEALK 167

RESULT 6
B64528
conserved hypothetical ATP-binding protein HP0066 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: B64528
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khatlak, H.G.; Glodok, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: B64528
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-831 <TOM>
A;Cross-references: UNIPROT:O24906; UNIPARC:UPI00000D308C; GB:AE000528; GB:AE000511; NI
C;Genetics:
A;Start codon: GTG
C;Keywords: nucleotide binding; P-loop
F;373-380/Region: nucleotide-binding motif A (P-loop)

Query Match 69.4%; Score 43; DB 2; Length 831;
 Best Local Similarity 69.2%; Pred. No. 49;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 |||:|||||
 Db 178 ILTESKEIEALK 190

RESULT 7
 H71038
 hypothetical protein PH1595 - *Pyrococcus horikoshii*
 C:Species: *Pyrococcus horikoshii*
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
 C:Accession: C71038
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: C71038
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-172 <KAW>
 A:Cross-references: UNIPROT:059254; UNIPARC:UPI00000630D5; GB:AP0000006; NID:g3236133; PI
 A:Experimental source: strain OT3
 A:Note: This accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH1595

Query Match 67.7%; Score 42; DB 2; Length 172;
 Best Local Similarity 69.2%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 |||||:||||
 Db 30 ILEENSSIINALK 42

RESULT 8
 H71647
 alanine-tRNA ligase (EC 6.1.1.7) (alas) RP856 - *Rickettsia prowazekii*
 C:Species: *Rickettsia prowazekii*
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
 C:Accession: H71647
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: H71647
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-877 <AND>
 A:Cross-references: UNIPROT:Q92C44; UNIPARC:UPI0000136321; GB:AJ235273; GB:AJ235269; NID
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: alas; RP856
 C:Superfamily: alanyl-tRNA ligase
 C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 67.7%; Score 42; DB 2; Length 877;
 Best Local Similarity 69.2%; Pred. No. 76;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 |||||:|||||:
 Db 738 ILERNKELEKELE 750

RESULT 9
 TSHUPL
 thrombospondin 1 precursor - human

C:Species: *Homo sapiens* (man)
 C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
 C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
 R:Lawler, J.; Hynes, R.O.
 J. Cell Biol. 103, 1635-1648, 1986
 A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c
 A:Reference number: A26155; MUID:87057617; PMID:2430973
 A:Accession: A26155
 A:Molecule type: mRNA
 A:Residues: 1-1170 <LAW>
 A:Cross-references: UNIPROT:P07996; UNIPARC:UPI0000046821; GB:X04665; NID:g37137; PIDN:C
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete
 R:Lahterty, C.D.; Gierman, T.M.; Dixit, V.M.
 J. Biol. Chem. 264, 11222-11227, 1989
 A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA s
 A:Reference number: A34274; MUID:89291870; PMID:2544587
 A:Accession: A34274
 A:Molecule type: DNA
 A:Residues: 1-166 <LAH>
 A:Cross-references: UNIPARC:UPI00001742BF; GB:J04835
 R:Hennessey, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,
 J. Cell Biol. 108, 729-736, 1989
 A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the
 A:Reference number: A30140; MUID:89139590; PMID:2918029
 A:Accession: A30140
 A:Molecule type: mRNA
 A:Residues: 1-83, 'A', '85-522, 'A', 524-1170 <HEN>
 A:Cross-references: UNIPARC:UPI0000038AB1; ENBL:X14787; NID:g37464; PIDN:CAA32889.1; PID
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete
 R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
 Biochemistry 25, 8418-8425, 1986
 A:Title: Partial amino acid sequence of human thrombospondin as determined by analysis o
 A:Reference number: A25812; MUID:87157592; PMID:3030396
 A:Accession: A25812
 A:Molecule type: mRNA
 A:Residues: 1-83, 'A', '85-397 <KOB>
 A:Cross-references: UNIPARC:UPI000016B0CA; GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:
 R:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
 A:Reference number: A05172; MUID:86287276; PMID:3461443
 A:Accession: A05172
 A:Molecule type: mRNA
 A:Residues: 1-83, 'A', '85-374, 'RC' <DIX>
 A:Cross-references: UNIPARC:UPI000016B140; GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete
 R:Sun, X.; Skorstengard, K.; Mosher, D.F.
 J. Cell Biol. 118, 693-701, 1992
 A:Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
 A:Reference number: A42927; MUID:92348511; PMID:1379247
 A:Accession: A42927
 A:Molecule type: protein
 A:Residues: 987-1003 <SUN>
 A:Cross-references: UNIPARC:UPI00001742C0
 A:Note: Cys-992 is shown to have a free sulphydryl
 C:Genetics:
 A:Gene: GDB:THBS1; TSP1; TSP
 A:Cross-references: GDB:120438; OMIM:188060
 A:Map position: 15q15-15q15
 A:Introns: 23/1
 A:Note: the list of introns may be incomplete
 C:Complex: homotrimer, disulfide linked
 C:Function:
 A:Description: participates in cell migration and adhesion, and in platelet aggregation
 A:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vc
 C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-1170/Product: thrombospondin 1 #status predicted <MAT>
 F:317-375/Domain: von Willebrand factor type C repeat homology <VNC>
 F:378-423/Domain: thrombospondin type 1 repeat homology <THR1>
 F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>
 F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>
 F:551-586/Domain: EGF homology <EGF1>
 F:650-689/Domain: EGF homology <EGF2>

F:926-928/Region: cell attachment (R-G-D) motif
F:171-232/Dsulfide bonds: #status predicted
F:248,376,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:270,274/Dsulfide bonds: interchain #status predicted
F:610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 67.7%; Score 42; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
DB 301 VTEENKELANEALR 313
: ||||| :
: ||||| :
: ||||| :

RESULT 10
T01241
probable MYB family transcription factor [imported] - Arabidopsis thaliana
N:Alternate names: myb-related transcription activator F16M14.2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Oct-2004
C:Accession: T01241; G84800
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, R.; Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, R. submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.
A:Reference number: Z14213
A:Accession: T01241
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-291 <ROU>
A:Cross-references: UNIPROT:O80435; UNIPARC:UPI00000A8813; EMBL:AC003028; NID:G3335356;
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84800
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-291 <STO>
A:Cross-references: UNIPARC:UPI00000A8813; GB:AE002093; NID:G3335378; PIDN:AAC27179.1; C
C:Genetics:
A:Gene: F16M14.2; At2g38090
A:Map position: 2
A:Introns: 142/2
F:14-68/Domain: myb DNA-binding repeat homology <MYB>

Query Match 66.1%; Score 41; DB 2; Length 291;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 EENKELENAL 12
DB 25 EENKFEENAL 34
: ||||| :
: ||||| :
: ||||| :

RESULT 11
AF1218
Salmonella typhimurium Pocr protein homolog lmo1150 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF1218
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AF1218
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <GLA>
A:Cross-references: UNIPROT:Q8Y7W9; UNIPARC:UPI0000054CCF; GB:NC_003210; PIDN:CAC99228.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1150

Query Match 66.1%; Score 41; DB 2; Length 294;
Best Local Similarity 72.7%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
DB 188 EENKEIRKALK 198
: ||||| :
: ||||| :
: ||||| :

RESULT 12
G72210
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: G72210
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of Thermotoga maritima.
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: G72210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <ARN>
A:Cross-references: UNIPROT:Q9X2C2; UNIPARC:UPI00000C11F6; GB:AE001817; GB:AE000512; NID:G3335356
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1802
C:Superfamily: Thermotoga maritima hypothetical protein TM1802

Query Match 66.1%; Score 41; DB 2; Length 546;
Best Local Similarity 61.5%; Pred. No. 66;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
DB 317 ILEESKIESQLK 329
: ||||| :
: ||||| :
: ||||| :

RESULT 13
A64675
alanine-CRNA ligase (EC 6.1.1.7) - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: A64675
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: A64675
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-847 <ROM>
A:Cross-references: UNIPROT:P56452; UNIPARC:UPI0000136308; GB:AE000629; GB:AE000511; NID:G3335356
C:Superfamily: alanyl-CRNA ligase
C:Keywords: ligase

Query Match 66.1%; Score 41; DB 2; Length 847;
Best Local Similarity 88.9%; Pred. No. 1.e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 BENKELENA 11
|||||:|
Db 699 BENKELKNA 707

RESULT 14

F71842
alanine-tRNA ligase (EC 6.1.1.7) - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: F71842
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: F71842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <ARN>
A:Cross-references: UNIPROT:Q9ZJY5; UNIPARC:UPI0000136307; GB:AE001543; GB:AE001439; NID
C:Experimental source: strain J99
C:Genetics:
A:Gene: alas
C:Superfamily: alanyl-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 66.1%; Score 41; DB 2; Length 847;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 BENKELENA 11
|||||:|
Db 699 EENKELKNA 707

RESULT 15

S7116
BCR-associated protein BAP31 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S7116
R:Adachi, T.; Schamel, W.W.A.; Kim, K.M.; Watanabe, T.; Becker, B.; Nielsen, P.J.; Reth,
EMBO J. 15, 1534-1541, 1996
A:Title: The specificity of association of the IGD molecule with the accessory proteins
A:Reference number: S7116; MUID:96203070; PMID:8612576
A:Accession: S7116
A:Molecule type: mRNA
A:Residues: 1-245 <ADA>
A:Cross-references: UNIPROT:Q61335; UNIPARC:UPI0000161CC5; EMBL:X81816; NID:G1487985; PI
C:Genetics:
A:Gene: BAP31
A:Map position: X

Query Match 64.5%; Score 40; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|||||:|
Db 177 LEENKSLKNDLR 189

RESULT 16

A53378
recombination protein recA - Thermus aquaticus
N:Alternate names: recombinase A
C:Species: Thermus aquaticus
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C:Accession: A53378; B55020
R:Angov, E.; Camerini-Otero, R.D.

J. Bacteriol. 176, 1405-1412, 1994
A:Title: The recA gene from the thermophile Thermus aquaticus YT-1: cloning, expression,
A:Reference number: A53378; MUID:94156846; PMID:8113181
A:Accession: A53378
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-340 <ANG>
A:Cross-references: UNIPROT:P48296; UNIPARC:UPI000013353D; GB:L20680; NID:G461416; PIDN:
A:Experimental source: strain YT-1
R:Wetmur, J.G.; Wong, D.M.; Ortiz, B.; Tong, J.; Reichert, F.; Gelfand, D.H.
J. Biol. Chem. 269, 25948-25935, 1994
A:Title: Cloning, sequencing, and expression of RecA proteins from three distantly relat.
A:Reference number: A55020; MUID:95014407; PMID:7929298
A:Accession: B55020
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-340 <WET>
A:Cross-references: UNIPARC:UPI000013353D; GB:L20095; NID:G349150; PIDN:AAA27502.1; PID:
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted

Query Match 64.5%; Score 40; DB 2; Length 340;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 2 LEEN--KELENALK 13
|||:|
Db 1 MEENKRKSLLENALK 14

RESULT 17

H72374
probable phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21) carbon dioxide-fixation
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: H72374
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: H72374
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <ARN>
A:Cross-references: UNIPROT:Q9WYS8; UNIPARC:UPI000000D3A00; GB:AE001723; GB:AE000512; NID:
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0447
C:Superfamily: phosphoribosylaminoimidazole carboxylase carbon dioxide-fixation chain; p
C:Keywords: carbon-carbon lyase; carboxy-lyase
F:15-349/Domain: phosphoribosylaminoimidazole carboxylase carbon dioxide-fixation chain

Query Match 64.5%; Score 40; DB 2; Length 380;
Best Local Similarity 53.8%; Pred. No. 66;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
||:|
Db 151 IIRKEDLENAIK 163

RESULT 18

G95085
conserved hypothetical protein SP0741 [imported] - Streptococcus pneumoniae (strain TIGR
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C:Accession: G95085
R:Tetellin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95085
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUR>
A:CROSS-references: UNIPROT:Q97RQ7; UNIPARC:UPI0000051578; GB:AE005672; PIDN:AAK74880.1;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0741

Query Match 64.5%; Score 40; DB 2; Length 391;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;
QY 2 LEENKE--LENALK 13
|||||:|||||
DB 269 LEENREKSLKALK 282

RESULT 19
C97953
conserved hypothetical protein spr0651 [imported] - *Streptococcus pneumoniae* (strain R6)
C:Species: *Streptococcus pneumoniae*
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: C97953
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: C97953
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <KUR>
A:CROSS-references: UNIPROT:Q8DQI7; UNIPARC:UPI00000E34C3; GB:AE007317; PIDN:AAK99455.1;
C:Genetics:
A:Gene: spr0651

Query Match 64.5%; Score 40; DB 2; Length 397;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;
QY 2 LEENKE--LENALK 13
|||||:|||||
DB 275 LEENREKSLKALK 288

RESULT 20
A55913
transcytosis-associated protein pl15 - rat
C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: A55913
R:Barroso, M.; Neilson, D.S.; Sztul, E.
Proc. Natl. Acad. Sci. U.S.A. 92, 527-531, 1995
A:Title: Transcytosis-associated protein (TAP)/pl15 is a general fusion factor required
A:Reference number: A55913; MUID:95132633; PMID:7831324
A:Accession: A55913
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-959 <BAR>
A:CROSS-references: UNIPROT:P41542; UNIPARC:UPI000016829C; GB:U15589; NID:G558474; PIDN:
Query Match 64.5%; Score 40; DB 2; Length 959;

Best Local Similarity 53.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
:||:||||:|
DB 853 LLQETKELKNEIK 865

RESULT 21
T43527
sp8 protein - fission yeast (*Schizosaccharomyces pombe*) (fragment)
C:Species: *Schizosaccharomyces pombe*
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 04-Feb-2000
C:Accession: T43527
R:Jimenez, M.; Goday, C.
submitted to the EMBL Data Library, June 1997
A:Description: *Schizosaccharomyces pombe* myosin-like.
A:Reference number: Z22545
A:Accession: T43527
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1173 <JIM>
A:CROSS-references: UNIPARC:UPI00001687F2; EMBL:AF010473; PIDN:AA65416.1
A:Experimental source: strain 972h(-)
C:Genetics:
A:Gene: sp8

Query Match 64.5%; Score 40; DB 2; Length 1173;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENAL 12
:||:|||||
DB 462 LLEENKHLNELL 473

RESULT 22
T50073
myosin-like coiled-coil protein sp8 [imported] - fission yeast (*Schizosaccharomyces pom*
C:Species: *Schizosaccharomyces pombe*
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50073
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.
submitted to the EMBL Data Library, December 1999
A:Reference number: Z25034
A:Accession: T50073
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1727 <MCD>
A:CROSS-references: UNIPROT:Q9UTK5; UNIPARC:UPI00001258AA; EMBL:AL133357; PIDN:CAB62414
A:Experimental source: strain 972h(-); cosmid c1486
C:Genetics:
A:Gene: SPDB:SPAC1486.04c
A:Map position: 1

Query Match 64.5%; Score 40; DB 2; Length 1727;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENAL 12
:||:|||||
DB 1016 LLEENKHLNELL 1027

RESULT 23
T12796
probable transglycosylase - *Bacillus subtilis* phage SPBc2
C:Species: *Bacillus subtilis* phage SPBc2
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T12796; A69911
R:Lazarevic, V.; Dueserhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A:Description: The complete nucleotide sequence of the *Bacillus subtilis* SPBc2 proph.

A;Reference number: Z17583
A;Accession: T12796
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2285 <LAZ>
A;Cross-references: UNIPROT:Q64046; UNIPARC:UPI00000605D0; EMBL:AF020713; NID:G3025478;
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortez
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
teich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69911
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2285 <KUN>
A;Cross-references: UNIPARC:UPI00000605D0; GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CA
A;Experimental source: strain 168
C;Genetics:
A;Gene: yomI

Query Match 64.5%; Score 40; DB 2; Length 2285;
Best Local Similarity 61.5%; Pred. No. 4+e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
||:|||||:
Db 847 ILKTNKELEKAIE 859

RESULT 24
T24864
hypothetical protein T12D8.1 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24864; T27892
R;McMurray, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19944
A;Accession: T24864
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2561 <W1L>
A;Cross-references: UNIPROT:Q46025; UNIPARC:UPI000007A66F; EMBL:Z81120; PIDN:CAB03348.1;
A;Experimental source: clone T12D8
R;Steward, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20436
A;Accession: T27892
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2561 <W12>
A;Cross-references: UNIPARC:UPI000007A66F; EMBL:Z82094; PIDN:CAB05024.1; GSPDB:GN00021;
A;Experimental source: clone ZK525
C;Genetics:
A;Gene: CESP-T12D8.1
A;Map position: 3
A;Introns: 44/3; 78/3; 104/2; 251/2; 466/3; 504/3; 548/3; 662/2; 747/1; 962/1; 1355/3; 1

Query Match 64.5%; Score 40; DB 2; Length 2561;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ENKELENAL 12
||| |||||:
Db 928 EENDELENSL 937

RESULT 25
H64034
hypothetical protein H11521 - *Haemophilus influenzae* (strain Rd KW20)
C;Species: *Haemophilus influenzae*
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: H64034
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.B.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: H64034
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-168 <TIGR>
A;Cross-references: UNIPROT:P44241; UNIPARC:UPI000013859F; GB:U32828; GB:L42023; NID:G15
C;Genetics:
A;Start codon: GTG

Query Match 62.9%; Score 39; DB 2; Length 168;
Best Local Similarity 59.2%; Pred. No. 40;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
| |||||:
Db 138 IFRNKELECLK 150

RESULT 26
E97074
Probable acetyltransferase [imported] - *Clostridium acetobutylicum*
C;Species: *Clostridium acetobutylicum*
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004
C;Accession: E97074
R;Nolling, J.; Berton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-171 <KUR>
A;Cross-references: UNIPROT:Q97J70; UNIPARC:UPI000000CA1C4; GB:AE001437; PIDN:AAK79384.1;
A;Experimental source: *Clostridium acetobutylicum* ATCC824
C;Genetics:
C;Superfamily: ribosomal protein serine N-acetyltransferase

Query Match 62.9%; Score 39; DB 2; Length 171;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
: ||| : |||||:
Db 43 VLEEREFIEAIAIK 55

RESULT 27
A86297
hypothetical protein F309.2 - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86297
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;

Mon Mar 13 10:54:38 2006

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712.
 A:Accession: A86297
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-230 <STO>
 A:Cross-references: UNIPROT:Q9SA21; UNIPARC:UPI00000A229D; GB:AE005172; NID:g4966342; P
 C:Genetics:
 A:Map position: 1

Query Match 62.9%; Score 39; DB 2; Length 230;
 Best Local Similarity 72.7%; Pred. No. 56;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 ENKELENALK 13
 : ||||| ||||
 Db 71 QEEKELEKALK 81

RESULT 28
 G86903
 cell shape determining protein [imported] - Lactococcus lactis subsp. lactis (strain IL1
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
 C:Accession: G86903
 R:Botolin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: G86903
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-291 <STO>
 A:Cross-references: UNIPROT:Q9CD19; UNIPARC:UPI00000C6BFE; GB:AE005176; PID:g12725299; P
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: mreC

Query Match 62.9%; Score 39; DB 2; Length 291;
 Best Local Similarity 80.0%; Pred. No. 72;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 ENKELENALK 13
 : ||||| ||||
 Db 102 ENKELEKALK 111

RESULT 29
 D64522
 carboxynorspermidine decarboxylase - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
 C:Accession: D64522
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.; McKenne
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: D64522
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-405 <TOM>
 A:Cross-references: UNIPROT:O24865; UNIPARC:UPI00000D2FC3; GB:AE000524; GB:AE000511; NID

Query Match 62.9%; Score 39; DB 2; Length 405;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELEN 10
 : ||||| ||||
 Db 121 ILDKNKQLEN 130

RESULT 30

H71984
 probable carboxynorspermidine decarboxylase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori
 A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: H71984

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: H71984

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <ARN>

A:Cross-references: UNIPROT:Q9ZM41; UNIPARC:UPI00000D35BA; GB:AE001441; GB:AE001439; NID

A:Experimental source: strain J99

C:Genetics:

A:Gene: nspC

Query Match 62.9%; Score 39; DB 2; Length 405;
 Best Local Similarity 70.0%; Pred. No. 1e+02;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELEN 10
 : ||||| ||||
 Db 121 ILDKNKQLEN 130

RESULT 31

B71083
 hypothetical protein PH0928 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C:Accession: B71083

R:Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: B71083

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-442 <KAM>

A:Cross-references: UNIPROT:O58688; UNIPARC:UPI0000062F42; GB:AP000004; NID:g3236131; P

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0928

Query Match 62.9%; Score 39; DB 2; Length 442;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
 : ||||| ||||
 Db 231 LEEGKSLEDAIR 242

RESULT 32

F75206

maltose-binding periplasmic protein precursor (maltodextrin-binding protein) (mmbp) PAB

C:Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F75206
R;anonymos, Genoscope
submitted to the ENBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A;Reference number: A75001
A;Accession: F75206
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <RAW>
A;Cross-references: UNIPROT:Q9V297; UNIPARC:UPI000012EB47; GB:AJ248283; GB:AL096836; NID
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: male; PAB0119

Query Match 62.9%; Score 39; DB 2; Length 453;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
|||:|||||:
Db 111 VFEQPDLENALK 123

RESULT 33
T02769
Gns protein - Adelaide River virus
C;Species: Adelaide River virus
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T02769
R;Wang, Y.; Walker, P.J.
Virology 195, 719-731, 1993
A;Title: Adelaide river rhabdovirus expresses consecutive glycoprotein genes as polycistronic
A;Reference number: Z14732; MUID:93331730; PMID:8337841
A;Accession: T02769
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: genomic RNA
A;Residues: 1-609 <RAW>
A;Cross-references: UNIPROT:Q89904; UNIPARC:UPI00000F4A64; EMBL:L09206; NID:g348027; PID
C;Superfamily: Adelaide River virus Gns protein

Query Match 62.9%; Score 39; DB 2; Length 609;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
|||:|||||:
Db 187 LESNKKLDNTVK 198

RESULT 34
A71023
arginine-tRNA ligase (EC 6.1.1.19) - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C;Accession: A71023
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: A71023
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-629 <RAW>
A;Cross-references: UNIPROT:Q59147; UNIPARC:UPI000013665C; GB:AP000006; NID:g3236133; PI
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1478
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 62.9%; Score 39; DB 2; Length 629;

Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
|||:|||||:
Db 218 LEDNPELENEIR 229

RESULT 35
AH2402
cell division protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2402
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2402
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-656 <KUR>
A;Cross-references: UNIPROT:Q8YMZ8; UNIPARC:UPI00000CEC4C; GB:BA000019; PIDN:BA076475.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: ftsH
C;Superfamily: cell division protein ftsH; FtsH/SEC18/Cdc48-type ATP-binding domain homo

Query Match 62.9%; Score 39; DB 2; Length 656;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
:|||||:
Db 118 LLEQNNELINRLK 130

RESULT 36
A81621
conserved hypothetical protein CP0020 [imported] - Chlamydomophila pneumoniae (strain AR39
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: A81621
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150285; PMID:10684935
A;Accession: A81621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-812 <REA>
A;Cross-references: UNIPROT:Q9K2E9; UNIPARC:UPI00000CCC37; GB:AE002166; GB:AE002161; NID
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0020

Query Match 62.9%; Score 39; DB 2; Length 812;
Best Local Similarity 72.7%; Pred. No. 2.1e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILEENKELENA 11
|||||:
Db 416 ILEEAKELDNS 426

RESULT 37
F72044
ct620 hypothetical protein - Chlamydomophila pneumoniae (strain CWL029)
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: F72044
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: F72044
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-831 <ARN>
A:Cross-references: UNIPROT:Q927H9; UNIPARC:UPI00000D3815; GB:AE001653; GB:AE001363; NID
A:Experimental source: strain CW1029
C:Genetics:
A:Gene: CPn0726

Query Match 62.9%; Score 39; DB 2; Length 831;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENA 11
||| |||:
Db 435 ILEAEKELDNS 445

RESULT 38
C86581
CT620 hypothetical protein [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86581
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: C86581
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-831 <STO>
A:Cross-references: UNIPROT:Q9JSB5; UNIPARC:UPI00000CCD1B; GB:BA000008; NID:g8979098; PI
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPj0726

Query Match 62.9%; Score 39; DB 2; Length 831;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENA 11
||| |||:
Db 435 ILEAEKELDNS 445

RESULT 39
G97865
alanine-tRNA ligase (EC 6.1.1.7) - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97865
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97865
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-878 <KUR>
A:Cross-references: UNIPROT:Q92G00; UNIPARC:UPI0000136320; GB:AE006914; PIDN:AAJ03865.1;
C:Genetics:
A:Gene: alas
C:Superfamily: alanyl-tRNA ligase
C:Keywords: ligase

Query Match 62.9%; Score 39; DB 2; Length 878;
Best Local Similarity 61.5%; Pred. No. 2.3e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENA 13
||| |||:
Db 738 ILEENKEVEKLE 750

RESULT 40
A99604
hypothetical protein MYPV 7370 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: A99604
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: A99604
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-991 <KUR>
A:Cross-references: UNIPROT:Q98PI5; UNIPARC:UPI00000C8105; GB:AL445566; PID:g14090152;
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 7370
A:Genetic code: SGC3

Query Match 62.9%; Score 39; DB 2; Length 991;
Best Local Similarity 53.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENA 13
||| |||:
Db 521 IVNENVEIENSLR 533

RESULT 41
I52300
giantin - human
N:Alternate names: gcp372
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: I52300
R:Soeda, M.; Misumi, Y.; Fujiwara, T.; Nishioka, M.; Ikehara, Y.
Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994
A:Title: Molecular cloning and sequence analysis of a human 372-kDa protein localized i
A:Reference number: I52300; MUID:95100974; PMID:7802676
A:Accession: I52300
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3225 <RES>
A:Cross-references: UNIPROT:Q14789; UNIPARC:UPI000016A987; GB:D25542; NID:g662389; PIDN
C:Superfamily: giantin

Query Match 62.9%; Score 39; DB 2; Length 3225;
Best Local Similarity 77.8%; Pred. No. 9.3e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LEENKELEN 10
||| |||:
Db 2525 LOQNKELN 2533

RESULT 42
A56539
giantin - human
N:Alternate names: macrogolin
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 26-Jan-1996 #text_change 09-Jul-2004
C:Accession: A56539; S37536
R:Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A:Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein

A;Reference number: A56539; MUID:94187728; PMID:7511208
A;Accession: A56539
A;Molecule type: mRNA
A;Residues: 1-3259 <SSE>
A;Cross-references: UNIPROT:Q14789; UNIPARC:UPI000012B42C; EMBL:X75304; NID:9405714; PID
C;Genetics:
A;Gene: GDB:GOLGB1; GCP; GCP371
A;Cross-references: GDB:454958
A;Map position: 3q13.31-3q13.31
C;Superfamily: giantin
C;Keywords: coiled coil; Golgi apparatus; transmembrane protein
P:3238-3254/Domain: transmembrane #status predicted <TMN>

Query Match 62.9%; Score 39; DB 1; Length 3259;
Best Local Similarity 77.8%; Pred. No. 9.4e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEENKELEN 10
Db 2559 LOONKELEN 2567
|:|||||

RESULT 43
H70333
conserved hypothetical protein aq_378 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: H70333
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V;
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: H70333
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-453 <AQF>
A;Cross-references: UNIPROT:Q66700; UNIPARC:UPI0000056337; GB:AE000687; NID:92983050; PI
A;Experimental source: strain VFS
C;Genetics:
A;Gene: aq_378

Query Match 62.1%; Score 38.5; DB 2; Length 453;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 3 EENKE-LENALK 13
Db 432 EQNKEKENALK 443
|:|||||

RESULT 44
A60457
clusterin - sheep (fragments)
N;Alternate names: SGP2; sulfated glycoprotein 2
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Oct-1995
C;Accession: A60457; S10671; S10672
R;Griswold, M.D.; Bishop, P.D.; Kim, K.H.; Ping, R.; Siiteri, J.E.; Morales, C.
Ann. N. Y. Acad. Sci. 564, 154-172, 1989
A;Title: Function of vitamin A in normal and synchronized seminiferous tubules.
A;Reference number: A60457; MUID:89372568; PMID:2672955
A;Accession: A60457
A;Molecule type: protein
A;Residues: 1-82 <GRI>
A;Cross-references: UNIPARC:UPI00001777B0
R;Tsututa, J.K.; Wong, K.; Fritz, I.B.; Griswold, M.D.
Biochem. J. 268, 571-578, 1990
A;Title: Structural analysis of sulphated glycoprotein 2 from amino acid sequence. Relat
A;Reference number: S10671; MUID:90303244; PMID:2363694
A;Accession: S10671
A;Molecule type: protein
A;Residues: 4-26;41-49 <TSU>

A;Cross-references: UNIPARC:UPI00001777B1; UNIPARC:UPI00001777B2
A;Accession: S10672
A;Molecule type: protein
A;Residues: 62-82 <TS2>
A;Cross-references: UNIPARC:UPI00001777B1; UNIPARC:UPI00001777B2
C;Superfamily: clusterin
C;Keywords: glycoprotein; testis

Query Match 61.3%; Score 38; DB 2; Length 82;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NKELENALK 13
Db 21 NKEIKNALK 29
|:|||||

RESULT 45
JN0729
hypothetical 16.2K protein - phage SPPI
N;Alternate names: hypothetical protein 42
C;Species: phage SPPI
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: JN0729; S43816; S41190; T42347; S21429
R;Chai, S.; Szepan, U.; Lueder, G.; Trautner, T.A.; Alonso, J.C.
Gene 129, 41-49, 1993
A;Title: Sequence analysis of the left end of the Bacillus subtilis bacteriophage SPPI
A;Reference number: JN0729; MUID:93328123; PMID:8335259
A;Accession: JN0729
A;Molecule type: DNA
A;Residues: 1-133 <CHA>
A;Cross-references: UNIPROT:Q38070; UNIPARC:UPI000009B0FE; EMBL:X65941; NID:914843; PIDN
R;Pedre, X.; Weise, F.; Chai, S.; Lueder, G.; Alonso, J.C.
J. Mol. Biol. 236, 1324-1340, 1994
A;Title: Analysis of cis and trans acting elements required for the initiation of DNA re
A;Reference number: S43798; MUID:94172631; PMID:8126723
A;Accession: S43816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <PED>
A;Cross-references: UNIPARC:UPI00001798BD; EMBL:X67865
R;Pedre, X.; Chai, S.; Lueder, G.; Trautner, T.A.; Alonso, J.C.
submitted to the EMBL Data Library, July 1992
A;Description: DNA replication studies with the Bacillus subtilis bacteriophage SPPI. Tt
A;Reference number: S41172
A;Accession: S41190
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 3-67 <PE2>
A;Cross-references: UNIPARC:UPI000016D74F; EMBL:X67865; NID:9472886; PIDN:CAA48065.1; P;J
R;Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A;Title: The complete nucleotide sequence and functional organization of Bacillus subtil
A;Reference number: 222137; MUID:98094274; PMID:9434185
A;Accession: T42347
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 3-133 <ALO>
A;Cross-references: UNIPARC:UPI00001637E1; EMBL:X97918; PIDN:CAA66501.1
C;Superfamily: phage SPPI hypothetical 16.2K protein

Query Match 61.3%; Score 38; DB 2; Length 133;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 EENKELENALK 13
Db 91 EENEDLKNELK 101
|:|||||

RESULT 46
A64516
hypothetical protein MJEC501 - Methanococcus jannaschii plasmid pURB801

C;Species: *Methanococcus jannaschii*
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
 C;Accession: A64516
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Iason, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
 A;Reference number: A64300, MUID:96337999; PMID:8688087
 A;Accession: A64516
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-276 <BUL>
 A;Cross-references: UNIPROT:Q60300; UNIPARC:UPI000013C255; GB:L77119; NID:gi500688; TIGR
 C;Genetics:
 A;Map position: ECSREV1259-429
 A;Genome: plasmid
 A;Note: this stable 16-kilobase pair plasmid is also designated ECS (small extrachromosomal)

Query Match 61.3%; Score 38; DB 2; Length 276;
 Best Local Similarity 72.7%; Pred. No. 98;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
 |||||
 DB 193 EYNKLENAPK 203

RESULT 47
 A11571
 Regulatory protein *Salmonella typhimurium* Pocr protein homolog lin114 [imported] - List
 C;Species: *Listeria innocua*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: A11571
 R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karet, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Munk, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative Genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: A11571
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-294 <GLA>
 A;Cross-references: UNIPROT:Q92CQ7; UNIPARC:UPI00000CC47D; GB:AL592022; PIDN:CAC96345.1
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: lin114

Query Match 61.3%; Score 38; DB 2; Length 294;
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
 |||||
 DB 188 EDNKEIRKALK 198

RESULT 48
 AB0078
 LysR-family regulatory protein YP06031 [imported] - *Yersinia pestis* (strain CO92)
 C;Species: *Yersinia pestis*
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AB0078
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
 A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AB0078
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-310 <KUR>
 A;Cross-references: UNIPROT:Q8Z180; UNIPARC:UPI00000DCE05; GB:AL590842; PIDN:CAC89485.1
 C;Genetics:
 A;Gene: YP06031
 C;Superfamily: conserved hypothetical protein H11364

Query Match 61.3%; Score 38; DB 2; Length 310;
 Best Local Similarity 69.2%; Pred. No. 1.1e+02;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 |||||
 DB 72 ILQEIDELEGALK 84

RESULT 49
 T33458
 Hypothetical protein F36H12.2 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T33458
 R;Blanchard, M.; Bradshaw, H.; Stellyes, L.
 submitted to the EMBL Data Library, July 1998
 A;Description: The sequence of C. elegans cosmid F36H12.
 A;Reference number: Z21346
 A;Accession: T33458
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-350 <BLA>
 A;Cross-references: UNIPROT:O76720; UNIPARC:UPI000007FBA7; EMBL:AF078790; PIDN:AAC26922
 A;Experimental source: strain Bristol N2; clone F36H12
 C;Genetics:
 A;Gene: CESP:F36H12.2
 A;Map position: 4
 A;Introns: 24/2; 172/3; 222/1; 312/2

Query Match 61.3%; Score 38; DB 2; Length 350;
 Best Local Similarity 61.5%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 |||||
 DB 294 ILRENHEIPNLIK 306

RESULT 50
 G71917
 8-amino-7-oxononanoate synthase (EC 2.3.1.47) - *Helicobacter pylori* (strain J99)
 C;Species: *Helicobacter pylori*
 A;Variety: strain J99
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
 C;Accession: G71917
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.
 A;Reference number: A71800; MUID:99120557; PMID:9923682
 A;Accession: G71917
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-373 <ARN>
 A;Cross-references: UNIPROT:Q9ZLN3; UNIPARC:UPI000012698A; GB:AE001487; GB:AE001439; NID:gi500688
 A;Experimental source: strain J99
 C;Genetics:
 A;Gene: bioF
 C;Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology
 C;Keywords: acyltransferase; coenzyme A

Query Match 61.3%; Score 38; DB 2; Length 373;
 Best Local Similarity 72.7%; Pred. No. 1.4e+02;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENAL 12
||:|||||

Db 350 LEDTKELANAL 360

RESULT 51

D72339

hypotheetical protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C:Accession: D72339

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of Thermotoga maritima

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: D72339

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-395 <ARN>

A:Cross-references: UNIPROT:Q9WZ36; UNIPARC:UPI00000C13A0; GB:AE001744; GB:AE000512; NID:10360571

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0737

C:Superfamily: Thermotoga maritima hypotheetical protein TM0737

Query Match 61.3%; Score 38; DB 2; Length 395;

Best Local Similarity 70.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 ENKELENALK 13
:||||:|||||

Db 109 KKKLQNALK 118

RESULT 52

D72324

O-acetylhomoserine sulphydrylase - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 05-Oct-2004

C:Accession: D72324

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of Thermotoga maritima

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: D72324

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-430 <ARN>

A:Cross-references: UNIPROT:Q9WZ4; UNIPARC:UPI00000D396B; GB:AE001753; GB:AE000512; NID:10360571

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0882

C:Superfamily: cystathionine gamma-synthase

Query Match 61.3%; Score 38; DB 2; Length 430;

Best Local Similarity 69.2%; Pred. No. 1.6e+02;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
| ||| |||||

Db 322 IAEGNKTRENALK 334

RESULT 53

A35744

clusterin precursor - bovine

N:Alternate names: complement cytolysis inhibitor; glycoprotein III

N:Contains: clusterin alpha chain; clusterin beta chain

C:Species: Bos primigenius taurus (cattle)

C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 09-Jul-2004

C:Accession: A35744

R:Palmer, D.J.; Christie, D.L.

J. Biol. Chem. 265, 6617-6623, 1990

A:Title: The primary structure of glycoprotein III from bovine adrenal medullary chromaffin cells

A:Reference number: A35744; MUID:90216681; PMID:1691174

A:Accession: A35744

A:Molecule type: mRNA

A:Residues: 1-439 <PAL>

A:Cross-references: UNIPROT:P17697; UNIPARC:UPI0000127BA3; GB:J05391; NID:gl63114; PIDN:10360571

A:Note: parts of this sequence, including the amino ends of the alpha and beta chains, were not sequenced

C:Superfamily: clusterin

C:Keywords: apoptosis; disulfide bond; extracellular protein; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-221/Domain: clusterin beta chain #status experimental <BCH>

F:222-439/Domain: clusterin alpha chain #status experimental <ACH>

F:80,97,139,283,320,346,366/Binding site: carbohydrate (Asn) #status predicted <AS>

Query Match 61.3%; Score 38; DB 2; Length 439;

Best Local Similarity 77.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 NKELENALK 13
|||:|||||

Db 37 NKEIKNALK 45

RESULT 54

A40018

clusterin precursor - dog

N:Alternate names: 80k secreted glycoprotein, renal, complement cytolysis inhibitor

C:Species: Canis lupus familiaris (dog)

C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004

C:Accession: A40018

R:Hartmann, K.; Rauch, J.; Urban, J.; Parczyk, K.; Diel, P.; Pilarsky, C.; Appel, D.; Haefliger, J.

J. Biol. Chem. 266, 9924-9931, 1991

A:Title: Molecular cloning of gp 80, a glycoprotein complex secreted by kidney cells in culture

A:Reference number: A40018; MUID:91236776; PMID:2033078

A:Accession: A40018

A:Molecule type: mRNA

A:Residues: 1-445 <HAR>

A:Cross-references: UNIPROT:P25473; UNIPARC:UPI0000127BA4; GB:M55251; GB:M38757; NID:gl63114; PIDN:10360571

A:Note: parts of this sequence, including the amino ends of the alpha and beta chains, were not sequenced

C:Superfamily: clusterin

C:Keywords: apoptosis; disulfide bond; glycoprotein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-226/Domain: clusterin beta chain #status experimental <BCH>

F:227-445/Domain: clusterin alpha chain #status experimental <ACH>

F:86,103,145,277,287,350,370/Binding site: carbohydrate (Asn) #status predicted <AS>

Query Match 61.3%; Score 38; DB 2; Length 445;

Best Local Similarity 77.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 NKELENALK 13
|||:|||||

Db 43 NKEIKNALK 51

RESULT 55

A42108

clusterin precursor - pig

N:Alternate names: complement cytolysis inhibitor; CP40 protein

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A42108; JCS535; FC4475

R:Diemer, V.; Hoyte, M.; Baglioni, C.; Millis, A.J.

J. Biol. Chem. 267, 5257-5264, 1992

A:Title: Expression of porcine complement cytolysis inhibitor mRNA in cultured aortic smooth muscle cells

A;Reference number: A42108; MUID:92184774; PMID:1544909

A;Accession: A42108
A;Molecule type: mRNA
A;Residues: 1-446 <DIE>
A;Cross-references: UNIPROT:Q29549; UNIPARC:UPI0000127BA8; GB:M84639; NID:g164408; PIDN:
A;Experimental source: aortic smooth muscle cells
A;Note: Sequence extracted from NCBI backbone (NCBIN:87354, NCBI:87356)
R;Ogawa, S.; Ishibashi, Y.; Sakamoto, Y.; Kitamura, K.; Kubo, M.; Sakai, T.; Inoue, K.
Biochem. Biophys. Res. Commun. 234, 712-718, 1997

A;Title: The glycoproteins that occur in the colloids of senescent porcine pituitary gla
A;Reference number: JCS535; MUID:97318844; PMID:9175781

A;Accession: JCS535

A;Molecule type: DNA

A;Residues: 1-446 <OGA>

A;Cross-references: UNIPARC:UPI0000127BA8

A;Accession: PC4475

A;Molecule type: protein

A;Residues: 58-66;68-77;229-247;249-251;408-436 <OG2>

A;Cross-references: UNIPARC:UPI0000127BA8

A;Experimental source: pituitary gland

C;Superfamily: clusterin

F;1-22/Domain: signal sequence #status predicted <SIG>

F;23-227/Domain: clusterin beta chain #status predicted <BCH>

F;23-227,228-446/Product: clusterin #status predicted <MAR>

F;228-446/Domain: clusterin alpha chain #status predicted <ACH>

Query Match 61.3%; Score 38; DB 2; Length 446;

Best Local Similarity 77.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 NKELENALK 13

|||||

Db 43 NKEIKNALK 51

RESULT 56

C81718

Signal recognition particle TC0294 [imported] - Chlamydia muridarum (strain Nigg)

C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: C81718

R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: C81718

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-448 <TET>

A;Cross-references: UNIPROT:Q9PL14; UNIPARC:UPI0000057886; GB:AE002297; GB:AE002160; NID

A;Experimental source: strain Nigg (MoPn)

C;Genetics:

A;Gene: TC0294

C;Superfamily: signal recognition particle 54K protein

Query Match 61.3%; Score 38; DB 2; Length 448;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENAL 12

|||||

Db 313 EENKELEKL 322

RESULT 57

D71566

Probable signal recognition particle GTPase - Chlamydia trachomatis (serotype D, strain

C;Species: Chlamydia trachomatis

C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004

C;Accession: D71566

R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998

A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra

A;Reference number: A71570; MUID:99000809; PMID:9784136

A;Accession: D71566

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-448 <ARN>

A;Cross-references: UNIPROT:O84028; UNIPARC:UPI00000D32BA; GB:AE001277; GB:AE001273; NID

A;Experimental source: serotype D, strain UW-3/Cx

C;Genetics:

A;Gene: ffh

C;Superfamily: signal recognition particle 54K protein

C;Keywords: nucleotide binding; P-loop

F;106-113/Region: nucleotide-binding motif A (P-loop)

Query Match 61.3%; Score 38; DB 2; Length 448;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENAL 12

|||||

Db 313 EENKELEKL 322

RESULT 58

T28878

hypothetical protein R08E3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T28878

R;Hawkins, J.

submitted to the EMBL Data Library, December 1995

A;Description: The sequence of C. elegans cosmid R08E3.

A;Reference number: Z20536

A;Accession: T28878

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-527 <RAW>

A;Cross-references: UNIPROT:Q21848; UNIPARC:UPI0000075A25; EMBL:U41027; PIDN:AAC48196.1

A;Experimental source: strain Bristol N2; clone R08E3

C;Genetics:

A;Gene: CESP.R08E3.3

A;Map position: X

A;Introns: 17/2; 38/2; 75/3; 118/1; 155/3; 244/3; 285/3; 323/3; 357/3; 426/3; 437/3; 48

Query Match 61.3%; Score 38; DB 2; Length 527;

Best Local Similarity 70.0%; Pred. No. 2e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 EENKELENAL 12

|||||

Db 254 EQNRELEKL 263

RESULT 59

S22594

hypothetical protein 1 - fruit fly (Drosophila subobscura) transposable element P

C;Species: Drosophila subobscura

C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: S22594

R;Paricio, N.; Perez-Alonso, M.; Martinez-Sebastian, M.J.; de Frutos, R.

Nucleic Acids Res. 19, 6713-6718, 1991

A;Title: P sequences of Drosophila subobscura lack exon 3 and may encode a 66 kd repres

A;Reference number: S22594; MUID:92107652; PMID:1662362

A;Accession: S22594

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-575 <PAR>

A;Cross-references: UNIPROT:Q24662; UNIPARC:UPI0000075F17; EMBL:X60436; NID:g11017; PID

C;Genetics:

A;Mobile element: transposable element P

A;Introns: 95/2; 322/1

Query Match 61.3%; Score 38; DB 2; Length 575;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
||:||||:|
DB 128 LEDYKELETSK 139

RESULT 60
S22595
hypothetical protein 2 - fruit fly (Drosophila subobscura) transposable element P
C:Species: Drosophila subobscura
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S22595
R:Paricio, N.; Perez-Alonso, M.; Martinez-Sebastian, M.J.; de Frutos, R.
Nucleic Acids Res. 19, 6713-6718, 1991
A:Title: P sequences of Drosophila subobscura lack exon 3 and may encode a 66 kd repress
A:Reference number: S22594; MUID:92107652; PMID:1662362
A:Accession: S22595
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-575 <PAR>
A:Cross-references: UNIPROT:Q24663; UNIPARC:UPI0000078C95; EMBL:X60436; NID:g11017; PID:
C:Genetics:
A:Mobile element: transposable element P
A:Introns: 95/2; 322/1

Query Match 61.3%; Score 38; DB 2; Length 575;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
||:||||:|
DB 128 LEDYKELETSK 139

Search completed: March 11, 2006, 12:25:52
Job time : 29.5263 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:12:17 ; Search time 127.263 Seconds
(without alignments)
72.070 Million cell updates/sec

Title: US-10-774-242A-6

Perfect score: 62

Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 70528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	760	1	SEPR HUMAN
2	62	100.0	760	2	Q53TP5 HUMAN
3	58	93.5	761	2	Q8R492 RAT
4	55	88.7	761	1	SEPR_MOUSE
5	50	80.6	352	2	Q5CDS3 CRYHO
6	50	80.6	1624	2	Q5CKS4 CRYPV
7	49	79.0	653	2	Q8H7V7 ORYSA
8	47	75.8	147	2	Q5XYQ4 BORGA
9	47	75.8	405	1	NIC1 SCHPO
10	45	72.6	162	2	Q97W88 SULSO
11	45	72.6	219	2	Q4XD55 PLACH
12	44	71.0	176	2	Q7WVB4 9STAP
13	44	71.0	759	2	Q5Z181 CHICK
14	44	71.0	2110	2	Q54BP1 DICDI
15	43	69.4	458	2	Q6N2T2 BRARE
16	43	69.4	469	2	Q4S9W6 TETNG
17	43	69.4	484	2	Q4S9L4 TETNG
18	43	69.4	532	2	Q7QF21 ANOGA
19	43	69.4	706	1	HDA1 YEAST
20	43	69.4	806	2	Q9ZW99 HELPY
21	43	69.4	831	2	Q24306 HELPY
22	42	67.7	172	2	Q59254 PYRHO
23	42	67.7	310	2	Q9PQF4 HEVER
24	42	67.7	310	2	Q66RN1 HEVER
25	42	67.7	342	2	Q8RFR9 FUSNN
26	42	67.7	459	2	Q86X02 HUMAN
27	42	67.7	523	2	Q4EG4 PLACH
28	42	67.7	534	2	Q6PNN2 HUMAN
29	42	67.7	535	2	Q15175 HUMAN
30	42	67.7	720	2	Q9H607 HUMAN
31	42	67.7	877	1	SYA_RICPR

32	42	67.7	877	1	SYA_RICTY	Q68VQ7 rickettsia
33	42	67.7	878	2	Q4UJTS_RICFE	Q4UJTS rickettsia
34	42	67.7	1170	1	TSP1_BOVIN	Q28178 bos taurus
35	42	67.7	1170	1	TSP1_HUMAN	P07996 homo sapien
36	42	67.7	1225	2	Q59E99 HUMAN	Q59E99 homo sapien
37	42	67.7	2479	2	Q66GS8 HUMAN	Q66GS8 homo sapien
38	41	66.1	73	2	Q8KZ79 BORBU	Q8KZ79 borrelia bu
39	41	66.1	126	2	Q735K3_BACCL	Q735K3 bacillus ce
40	41	66.1	140	2	Q8BZ17 SCHJA	Q8BZ17 schistosoma
41	41	66.1	211	2	Q81337 PLAF7	Q81337 plasmodium
42	41	66.1	258	2	Q4L9M7 BORBU	Q4L9M7 borrelia bu
43	41	66.1	258	2	Q950J2 BORBU	Q950J2 borrelia bu
44	41	66.1	258	2	Q80435 ARATH	Q80435 arabidopsis
45	41	66.1	294	2	Q720T1 LISMF	Q720T1 listeria mo
46	41	66.1	294	2	Q8Y7W9 LISMO	Q8Y7W9 listeria mo
47	41	66.1	298	2	Q8GXN7 ARATH	Q8GXN7 arabidopsis
48	41	66.1	298	2	Q8KU29 CLODI	Q8KU29 clostridium
49	41	66.1	395	2	Q58EL9 BRARE	Q58EL9 brachydanio
50	41	66.1	420	2	Q6AEY7_LEIXX	Q6AEY7 leifsonia x
51	41	66.1	440	2	Q8U3F5 PYRFU	Q8U3F5 pyrococcus
52	41	66.1	459	2	Q4Y6Y2 PLACH	Q4Y6Y2 plasmodium
53	41	66.1	496	2	Q9X2C2_THEMA	Q9X2C2 thermotoga
54	41	66.1	546	2	Q62319_CAEEL	Q62319 caenorhabdi
55	41	66.1	606	2	Q8KXW2 CLOSTR	Q8KXW2 clostridium
56	41	66.1	610	2	Q8KXW2 CLOSTR	Q8KXW2 clostridium
57	41	66.1	784	2	Q54CU5 DICDI	Q54CU5 dictyosteli
58	41	66.1	790	2	Q9VIK6 DROME	Q9VIK6 drosophila
59	41	66.1	847	1	SYA_HELPR	Q92JY5 helicobacte
60	41	66.1	847	1	SYA_HELPR	P56452 helicobacte
61	41	66.1	877	2	Q55E10 DICDI	Q55E10 dictyosteli
62	41	66.1	935	2	Q54P19 DICDI	Q54P19 dictyosteli
63	41	66.1	951	2	Q7RKA6 PLAMO	Q7RKA6 plasmodium
64	41	66.1	1027	2	Q7RCL2 PLAMO	Q7RCL2 plasmodium
65	41	66.1	1037	2	Q5CW77 CRYPV	Q5CW77 cryptospori
66	41	66.1	1037	2	Q7YXN7 CRYPV	Q7YXN7 cryptospori
67	41	66.1	1040	2	Q5CIN0_CRYHO	Q5CIN0 cryptospori
68	41	66.1	1076	2	Q4Z406 PLABE	Q4Z406 plasmodium
69	41	66.1	1135	2	Q68XX2 RICTY	Q68XX2 rickettsia
70	41	66.1	1632	2	Q8IKM7 PLAF7	Q8IKM7 plasmodium
71	41	66.1	2193	2	Q4PFN8 USTMA	Q4PFN8 ustilago ma
72	40	64.5	68	2	Q5XZ82 BORGA	Q5XZ82 borrelia ga
73	40	64.5	99	2	Q4FU05_9GAMM	Q4FU05 psychrobact
74	40	64.5	122	2	Q5XYE9 BORGA	Q5XYE9 xenopus tro
75	40	64.5	202	2	Q5EAM0_XENTR	Q5EAM0 xenopus tro
76	40	64.5	229	2	Q5CE31_CRYHO	Q5CE31 cryptospori
77	40	64.5	244	1	BAP31_MOUSE	Q61335 mus musculu
78	40	64.5	246	2	Q8RD89_THETN	Q8RD89 thermotoga
79	40	64.5	247	1	SCPA_THETN	Q8RAA1 thermotoga
80	40	64.5	276	2	Q9RQ94_RHLET	Q9RQ94 rhizobium e
81	40	64.5	335	2	Q99JZ5_MOUSE	Q99JZ5 mus musculu
82	40	64.5	340	1	RECA_THEAO	P48296 thermus aqu
83	40	64.5	354	2	Q9EN24_AMRPV	Q9EN24 amsaeta moo
84	40	64.5	380	2	Q9WYS8_THEMA	Q9WYS8 thermotoga
85	40	64.5	391	2	Q4MZT2_THSPA	Q4MZT2 theileria p
86	40	64.5	391	2	Q97RQ7_STRPN	Q97RQ7 streptococc
87	40	64.5	397	2	Q8DQ17_STRRP	Q8DQ17 streptococc
88	40	64.5	456	2	Q84472 DICDI	Q84472 dictyosteli
89	40	64.5	496	2	Q84VF9_ORYSA	Q84VF9 oryza sativ
90	40	64.5	522	2	Q54Q68 DICDI	Q54Q68 dictyosteli
91	40	64.5	624	2	Q891B9_CLOTE	Q891B9 clostridium
92	40	64.5	646	2	Q4JMX3_9BACT	Q4JMX3 uncultured
93	40	64.5	678	2	Q5R7N7_PONPY	Q5R7N7 pongo pygma
94	40	64.5	711	2	Q9LRM8_ARATH	Q9LRM8 arabidopsis
95	40	64.5	767	2	Q6C919_YARLI	Q6C919 yarrowia li
96	40	64.5	795	2	Q7F099_ORYSA	Q7F099 oryza sativ
97	40	64.5	833	2	Q6CFK7_YARLI	Q6CFK7 yarrowia li
98	40	64.5	917	2	Q9LKU3_ARATH	Q9LKU3 arabidopsis
99	40	64.5	931	2	Q8L1B1_ORYSA	Q8L1B1 oryza sativ
100	40	64.5	941	1	VDP_MOUSE	Q92120 mus musculu
101	40	64.5	959	2	Q91WE7_MOUSE	Q91WE7 mus musculu
102	40	64.5	961	1	VDP_BOVIN	P41541 bos taurus
103	40	64.5	961	1	VDP_BOVIN	P41541 bos taurus
104	40	64.5	961	2	Q8N592_HUMAN	Q8N592 homo sapien

105	40	64.5	962	1	VDP_HUMAN	O60763	homo sapien	178	39	62.9	629	1	SYR_PYRHO	O59147	pyrococcus
106	40	64.5	973	2	O65T88_HUMAN	O86tb8	homo sapien	179	39	62.9	648	2	O7KWN1_DICDI	O7kwn1	dictyosteli
107	40	64.5	990	2	O9C799_ATH	O9c793	arabidopsis	180	39	62.9	656	2	O8Ym28_ANASP	O8ym28	anabaena sp
108	40	64.5	1151	2	O51H43_ENTHI	O51h43	entamoeba h	181	39	62.9	663	2	O8TMY1_METAC	O8tmy1	methanocarc
109	40	64.5	1249	2	O812R2_PLAF7	O812r2	plasmodium	182	39	62.9	672	2	O55AJ7_DICDI	O55aj7	dictyosteli
110	40	64.5	1335	2	O9GP66_DROME	O9gp66	drosophila	183	39	62.9	772	2	O75J95_DICDI	O75j95	dictyosteli
111	40	64.5	1335	2	O9W1X5_DROME	O9w1x5	drosophila	184	39	62.9	797	1	TAMO_DROME	O9wia4	drosophila
112	40	64.5	1328	2	O6CKA9_KLULA	O6cka9	klyveromyc	185	39	62.9	804	2	O5L1L0_GROKA	O5l1l0	geobacillus
113	40	64.5	1277	1	ALMI_SCHPO	O9uck5	schizosach	186	39	62.9	812	2	O9K2B9_CHLPN	O9k2b9	chlamydia p
114	40	64.5	1864	2	O6LFE6_PLAF7	O6lfe6	plasmodium	187	39	62.9	831	2	O7VPV8_CHLPN	O7vpv8	chlamydia p
115	40	64.5	2285	2	O64046_BPSPC	O64046	bacterioph	188	39	62.9	831	2	O9JBS5_CHLPN	O9jbs5	chlamydia p
116	40	64.5	2285	2	O31976_BACSU	O31976	bacillus su	189	39	62.9	831	2	O927H9_CHLPN	O927h9	chlamydia p
117	40	64.5	2561	2	O46025_CABEL	O46025	caenorhabdi	190	39	62.9	837	2	O4XQB4_PLACH	O4xqe4	plasmodium
118	40	64.5	4454	2	O4FYE6_LEIMA	O4fye6	leishmania	191	39	62.9	856	2	O4HM28_CAMLA	O4hxm8	campylobact
119	40	64.5	4470	2	O4KXA7_LEIMA	O4kxa7	leishmania	192	39	62.9	868	2	O4YXD3_PLABE	O4ymd3	plasmodium
120	39.5	63.7	288	2	O4HSS5_CAMUP	O4hs55	campylobact	193	39	62.9	878	1	SYA_RICGN	O92g00	rickettsia
121	39	62.9	77	1	YIBR_SALTY	O4hs85	campylobact	194	39	62.9	878	2	O7PA66_RICSI	O7pa66	rickettsia
122	39	62.9	77	1	YIBR_SALTY	P40824	salmonella	194	39	62.9	878	2	O7PA66_RICSI	O7pa66	rickettsia
123	39	62.9	77	1	YIBR_SALTY	O57ib3	salmonella	195	39	62.9	982	2	O8IDQ6_PLAF7	O8ldq6	plasmodium
124	39	62.9	98	2	O4FR22_9GAMM	O4fr22	psychrobact	196	39	62.9	991	2	O98P15_MYCPU	O98p15	mycoplasma
125	39	62.9	130	2	O8XL80_CLOPE	O8xl80	clostridium	197	39	62.9	1009	2	O5YFD7_9VIRU	O5yfd7	singapore g
126	39	62.9	145	2	O4X8J8_PLACH	O4x8j8	plasmodium	198	39	62.9	1009	2	O5GAE5_9VIRU	O5gae5	grouper iri
127	39	62.9	168	1	VG8_HABIN	P42441	haemophilus	199	39	62.9	1058	2	O4S9X2_TETNG	O4s9x2	tetraodon n
128	39	62.9	171	2	O97J70_CLOAB	O97j70	clostridium	200	39	62.9	1059	2	O7RK18_PLAYO	O7rk18	plasmodium
129	39	62.9	203	2	O8FNK0_GRIKC	O8fnk0	caudatus	201	39	62.9	1098	2	O4W193_ASFPF	O4wi93	aspergillus
130	39	62.9	230	2	O9SA21_ATH	O9sa21	arabidopsis	202	39	62.9	1245	2	O7V1J2_PROMP	O7v1j2	prochloroco
131	39	62.9	233	2	O93C53_ENTFA	O93c53	enterococcu	203	39	62.9	1250	2	O4Z5G2_PLABE	O4z5g2	plasmodium
132	39	62.9	234	2	O8YVC9_ATH	O8yvc9	arabidopsis	204	39	62.9	1333	2	O7QCA4_ANOGA	O7qca4	anopheles g
133	39	62.9	246	2	O72QR6_LEPIC	O72qr6	leptospira	206	39	62.9	1333	2	O7QCA4_ANOGA	O7qca4	anopheles g
134	39	62.9	261	2	O51AP9_ENTHI	O51ap9	entamoeba h	207	39	62.9	1770	2	O8FT93_CANGA	O8ft93	candida gla
135	39	62.9	274	2	O5RCC3_PONPY	O5rc33	pongo pygma	208	39	62.9	1818	2	O8I157_PLAF7	O8i157	plasmodium
136	39	62.9	283	2	O51CQ7_ENTHI	O51cq7	entamoeba h	209	38.5	62.1	3259	1	O6GB1_HUMAN	O6gb1	homo sapien
137	39	62.9	291	2	O9CDI9_LACLA	O9cdi9	lactococcus	210	38.5	62.1	453	2	O66700_AQUAE	O66700	aquifex aeo
138	39	62.9	308	2	O4HNH4_CAMUP	O4hnh4	campylobact	211	38	61.3	1034	2	O5AHD2_CANAL	O5ahd2	candida alb
139	39	62.9	314	1	NMI_MOUSE	O35309	mus musculus	212	38	61.3	66	1	CLUS_SHEEP	P17698	ovis aries
140	39	62.9	314	2	O4FJP2_MOUSE	O4fjp2	mus musculus	213	38	61.3	86	2	O5KY80_GROKA	O5ky80	geobacillus
141	39	62.9	318	2	O81JT6_PLAF7	O81jt6	plasmodium	214	38	61.3	94	2	O9X3N8_BORBU	O9x3n8	borrelia bu
142	39	62.9	320	2	O7U8H6_SYNPX	O7u8h6	synecococc	215	38	61.3	94	2	O5L2Q0_STRT1	O5l2q0	streptococc
143	39	62.9	321	2	O54B11_DICDI	O54b11	dictyosteli	216	38	61.3	94	2	O5M4B5_STRT2	O5m4b5	streptococc
144	39	62.9	346	2	O6V212_CNPV	O6v212	canarypox v	217	38	61.3	102	2	O9R9B9_BORBU	O9r9b9	borrelia bu
145	39	62.9	376	2	O5LEB1_BACFN	O5leb1	bacteroides	218	38	61.3	130	2	O723Q0_LISMF	O723q0	liesteria mo
146	39	62.9	387	2	O96WH9_ZYGRO	O96wh9	zygosacchar	219	38	61.3	131	2	O7YQF2_BSPPP	O7yqf2	bacterioph
147	39	62.9	401	2	O4N549_THERPA	O4n549	theliera p	220	38	61.3	131	2	O38070_BPSPP	O38070	bacterioph
148	39	62.9	405	2	O24865_HELPY	O24865	helicobacte	221	38	61.3	138	2	O66Z87_BORBU	O66z87	borrelia bu
149	39	62.9	405	2	O9ZN41_HELPFJ	O9zn41	helicobacte	222	38	61.3	146	2	O4X912_PLACH	O4x912	plasmodium
150	39	62.9	410	2	O5TZA5_BRARE	O5tza5	brachydanio	223	38	61.3	155	2	O4Z901_9CAUD	O4z901	staphylococ
151	39	62.9	410	2	O803P8_BRARE	O803p8	brachydanio	224	38	61.3	196	2	O86118_BORBU	O86118	borrelia bu
152	39	62.9	411	2	O8VC91_MOUSE	O8vc91	mus musculus	225	38	61.3	199	2	O86EV9_SCHJA	O86ev9	schistosoma
153	39	62.9	412	1	NEMO_MOUSE	O88322	m nf-kappa-	226	38	61.3	221	2	O86VERO_MOUSE	O86ver0	mus musculus
154	39	62.9	412	2	O6TNG5_RAT	O6tng5	rattus norv	227	38	61.3	225	2	O66Z80_BORBU	O66z80	borrelia bu
155	39	62.9	412	2	O924H4_MOUSE	O924h4	mus musculus	228	38	61.3	225	2	O66Z80_BORBU	O66z80	borrelia bu
156	39	62.9	418	2	O5E9T0_BOVIN	O5e9t0	bos taurus	229	38	61.3	231	2	O4RYL4_TETNG	O4ryl4	tetraodon n
157	39	62.9	419	1	NEMO_HUMAN	O9yck9	h nf-kappa-	230	38	61.3	231	2	O61KU8_CABER	O61ku8	caenorhabdi
158	39	62.9	419	2	O7LBV6_HUMAN	O7lbv6	homo sapien	231	38	61.3	255	2	O66Z89_BORBU	O66z89	borrelia bu
159	39	62.9	419	2	O95KU9_BOVIN	O95ku9	bos taurus	232	38	61.3	256	2	O66Z95_BORBU	O66z95	borrelia bu
160	39	62.9	419	2	O5EA16_BOVIN	O5ea16	bos taurus	233	38	61.3	257	2	O9S0A8_BORBU	O9s0a8	borrelia bu
161	39	62.9	425	1	CNI130_HUMAN	O8ns06	homo sapien	234	38	61.3	257	2	O66Z71_BORBU	O66z71	borrelia bu
162	39	62.9	425	1	CNI130_MOUSE	O8nb04	mus musculus	235	38	61.3	257	2	O66Z84_BORBU	O66z84	borrelia bu
163	39	62.9	425	2	O4R222_MACFA	O4r222	macaca fasc	236	38	61.3	258	2	O66Z92_BORBU	O66z92	borrelia bu
164	39	62.9	425	2	O642A8_RAT	O642a8	rattus norv	237	38	61.3	258	2	O66Z73_BORBU	O66z73	borrelia bu
165	39	62.9	425	2	O52K00_MOUSE	O52k00	mus musculus	238	38	61.3	258	2	O87304_BORBU	O87304	borrelia bu
166	39	62.9	430	2	O7TSS3_MOUSE	O7tss3	mus musculus	239	38	61.3	258	2	O44753_BORBU	O44753	borrelia bu
167	39	62.9	442	2	O58688_PYRHO	O58688	pyrococcus	240	38	61.3	258	2	O44789_BORBU	O44789	borrelia bu
168	39	62.9	453	1	MALE_PYRAB	O9v297	pyrococcus	241	38	61.3	258	2	O5XYM2_BORGA	O5xym2	borrelia ga
169	39	62.9	487	2	O7Z7F1_HUMAN	O7z7f1	homo sapien	242	38	61.3	258	2	O4Y7M2_BORBU	O4y7m2	borrelia bu
170	39	62.9	487	2	O4RLJ6_TETNG	O4rlj6	tetraodon n	243	38	61.3	263	2	O86122_BORBU	O86122	borrelia bu
171	39	62.9	499	2	O7RLD6_PLAYO	O7rld6	plasmodium	244	38	61.3	263	2	O4XDA4_PLACH	O4xda4	plasmodium
172	39	62.9	515	2	O57Z13_9TRYP	O57z13	trypanosoma	245	38	61.3	276	1	Y3401_METJA	Y3401	methanococc
173	39	62.9	536	2	O5C9P0_CRYPV	O5c9p0	cryptospori	246	38	61.3	292	2	O9L632_MYCHO	O9l632	mycoplasma
174	39	62.9	546	2	O7QA89_ANOGA	O7qa89	anopheles g	247	38	61.3	294	2	O92CQ7_LISIN	O92cq7	liesteria in
175	39	62.9	595	2	O6BU56_DEBHA	O6bu56	debaromyces	248	38	61.3	304	2	O91FB6_IRV6	O91fb6	chilo iride
176	39	62.9	602	2	O813C2_BACCR	O813c2	bacillus ce	249	38	61.3	308	2	O8VGE4_MOUSE	O8vge4	mus musculus
177	39	62.9	609	2	O89304_9RHAB	O89304	adelaide ri	250	38	61.3	310	2	O8Z180_YERPE	O8z180	yersinia pe

251	38	61.3	312	2	Q7TQX5_MOUSE	Q7tqx5 mus musculus	324	38	61.3	841	2	Q7SAXO_NEUCR	Q7sax0 neurospora
252	38	61.3	313	2	Q5E386_VIBF1	Q5e386 vibrio fisc	325	38	61.3	843	1	COT_HUMAN	C1643 homo sapien
253	38	61.3	314	2	Q83CN4_COXBU	Q83cn4 coxiella bu	326	38	61.3	843	2	QSRADO_PONPY	Qsrado pongo pygma
254	38	61.3	315	2	Q7TS25_MOUSE	Q7ts25 mus musculus	327	38	61.3	860	2	Q4RQJ4_TETNG	Q4rqj4 tetraodon n
255	38	61.3	315	2	Q7TS28_MOUSE	Q7ts28 mus musculus	328	38	61.3	895	2	Q7SHU8_NEUCR	Q7shu8 neurospora
256	38	61.3	318	2	Q665T7_YERPS	Q665t7 yersinia ps	329	38	61.3	910	2	Q7WRX0_STAAU	Q7wrx0 staphylococ
257	38	61.3	318	2	Q8CG22_MOUSE	Q8cg22 mus musculus	330	38	61.3	932	2	Q9SU74_STAAU	Q9su74 staphylococ
258	38	61.3	326	2	Q6MDV3_PARUM	Q6mdv3 parachlamyd	331	38	61.3	932	2	Q7ASB7_STAAU	Q7asb7 staphylococ
259	38	61.3	328	2	Q7TRL9_MOUSE	Q7trl9 mus musculus	332	38	61.3	932	2	Q7WRM3_STAAU	Q7wrm3 staphylococ
260	38	61.3	328	2	Q8VEV4_MOUSE	Q8vev4 mus musculus	333	38	61.3	932	2	Q6G9E8_STAAU	Q6g9e8 staphylococ
261	38	61.3	335	2	Q877G1_SULAC	Q877g1 sulfolobus	334	38	61.3	932	2	Q6GGZ5_STAAU	Q6ggz5 staphylococ
262	38	61.3	342	2	Q7P6P3_FUSNV	Q7p6p3 fusobacteri	335	38	61.3	932	2	Q5HG06_STAAU	Q5hg06 staphylococ
263	38	61.3	347	2	Q6R7D5_9HERP	Q6r7d5 ostreid her	336	38	61.3	932	2	Q8NWR6_STAAU	Q8nwr6 staphylococ
264	38	61.3	352	2	Q76720_CABEL	Q76720 caenorhabdi	337	38	61.3	932	2	Q931R8_STAAU	Q931r8 staphylococ
265	38	61.3	357	2	Q4Y082_PLACH	Q4y082 plasmodium	338	38	61.3	977	2	Q5CVH7_CRYPV	Q5cvh7 cryptospori
266	38	61.3	359	2	Q6VZC0_CNPV	Q6vzc0 canarypox v	339	38	61.3	998	2	Q59FY4_HUMAN	Q59fy4 homo sapien
267	38	61.3	368	1	MNN9_CANAL	P53697 candida alb	340	38	61.3	1014	2	Q8RG94_FUSNN	Q8rg94 fusobacteri
268	38	61.3	368	2	Q5A4F1_CANAL	Q5a4f1 candida alb	341	38	61.3	1096	2	Q54DR3_DICDI	Q54dr3 dictyosteli
269	38	61.3	371	2	Q6VZB9_CNPV	Q6vzb9 canarypox v	342	38	61.3	1111	2	Q9UOK5_PLAF7	Q9uok5 plasmodium
270	38	61.3	373	1	BIOF_HELPJ	Q92ln3 helicobacte	343	38	61.3	1142	2	Q54F44_DICDI	Q54f44 dictyosteli
271	38	61.3	377	2	Q5NF08_FRATT	Q5nf08 francisella	344	38	61.3	1142	2	Q8GMX0_DICDI	Q8gmx0 dictyosteli
272	38	61.3	377	2	Q5C8D3_9PEZI	Q5c8d3 chaetophae	345	38	61.3	1154	2	Q7PTG9_ANOGA	Q7ptg9 anopheles g
273	38	61.3	395	2	Q9WZJ6_THENA	Q9wzj6 thermotoga	346	38	61.3	1235	2	Q8SX64_DROME	Q8sx64 drosophila
274	38	61.3	396	2	Q6VZ17_CNPV	Q6vz17 canarypox v	347	38	61.3	1320	2	Q9JK25_RAT	Q9jk25 rattus norv
275	38	61.3	409	2	Q9BLX6_CIOIN	Q9blx6 ciona intes	348	38	61.3	1381	2	Q9VZT7_DROME	Q9vzt7 drosophila
276	38	61.3	430	2	Q7PDK9_PLAYO	Q7pdk9 plasmodium	349	38	61.3	1405	2	Q8EVL6_MYCPE	Q8evl6 mycoplasma
277	38	61.3	430	2	Q9WZY4_THEMA	Q9wzy4 thermotoga	350	38	61.3	1463	2	Q7QJR3_ANOGA	Q7qjr3 anopheles g
278	38	61.3	435	2	Q4YMS7_PLABE	Q4ym57 plasmodium	351	38	61.3	1464	2	Q7SHT9_NEUCR	Q7sht9 neurospora
279	38	61.3	437	2	Q4XWM1_ASPPU	Q4xwm1 aspergillus	352	38	61.3	1489	2	Q6OSE8_CABER	Q6ose8 caenorhabdi
280	38	61.3	437	2	Q4XWF2_PLACH	Q4xwf2 plasmodium	353	38	61.3	1508	2	Q4WNH0_ASPPU	Q4wnh0 aspergillus
281	38	61.3	439	1	CLUS_BOVIN	P17697 bos taurus	354	38	61.3	1521	2	Q7UZY3_PROMP	Q7uzy3 prochloroco
282	38	61.3	445	1	CLUS_CANFA	P25473 canis famil	355	38	61.3	1527	2	Q7RIN9_PLAYO	Q7rin9 plasmodie
283	38	61.3	446	1	CLUS_PIG	Q29549 sus scrofa	356	38	61.3	1550	2	Q525D1_MAGGR	Q525d1 magnaporthe
284	38	61.3	448	2	Q84028_CHLTR	Q84028 chlamydia t	357	38	61.3	1558	2	Q55VU3_CRYNE	Q55vu3 cryptococcu
285	38	61.3	448	2	Q9PL14_CHLMU	Q9pl14 chlamydia t	358	38	61.3	1558	2	Q5KKE6_CRYNE	Q5kke6 cryptococcu
286	38	61.3	449	1	CLUS_HORSE	Q29482 equus cabal	359	38	61.3	1862	2	Q5L1P5_MAGGR	Q5l1p5 magnaporthe
287	38	61.3	453	2	Q4YTA8_PLABE	Q4yta8 plasmodium	360	38	61.3	1915	2	Q5COE6_CRYPV	Q5coe6 cryptospori
288	38	61.3	455	2	Q21848_CABEL	Q21848 caenorhabdi	361	38	61.3	2346	2	Q8UYV9_9VIRU	Q8uyv9 strawberry
289	38	61.3	458	2	Q4XFN9_PLACH	Q4xfn9 plasmodium	362	38	61.3	2346	2	Q7Z561_HUMAN	Q7z561 homo sapien
290	38	61.3	462	2	Q5CKZ7_CRYHO	Q5ckz7 cryptospori	363	38	61.3	2346	2	Q6XDA8_HUMAN	Q6xda8 homo sapien
291	38	61.3	475	2	Q74MT9_NANPQ	Q74mt9 nanoarchaeu	364	38	61.3	2360	2	Q4INY7_GIBZE	Q4iny7 gibberella
292	38	61.3	480	2	Q4YUW3_PLABE	Q4yuw3 plasmodium	365	38	61.3	2369	2	Q8IBV8_PLAF7	Q8ibv8 plasmodium
293	38	61.3	480	2	Q9DMV2_MOUSE	Q9dmv2 mus musculus	366	38	61.3	279	2	Q7N3P5_PHOLL	Q7n3p5 photorhabdu
294	38	61.3	480	2	Q9D9F0_MOUSE	Q9d9f0 mus musculus	367	38	61.3	279	2	Q5HVD5_CAMJR	Q5hvd5 campylobact
295	38	61.3	486	2	Q7RKU4_PLAYO	Q7rku4 plasmodium	368	38	61.3	279	2	Q4HHN1_CAMJE	Q4hnh1 campylobact
296	38	61.3	516	2	Q7JP60_CABEL	Q7jp60 caenorhabdi	369	38	61.3	291	2	Q4HHW1_CAMCO	Q4hwh1 campylobact
297	38	61.3	517	1	OPTN_BRARE	Q5r156 brachydanio	370	38	61.3	583	2	Q41869_STAHJ	Q41869 staphylococ
298	38	61.3	530	2	Q7VG94_HELHP	Q7vg94 helicobacte	371	38	61.3	583	2	Q4RAA6_TETNG	Q4raa6 tetraodon n
299	38	61.3	534	2	Q4KXG6_PSEF5	Q4kxc6 pseudomonas	372	38	61.3	583	2	Q4RAA5_TETNG	Q4raa5 tetraodon n
300	38	61.3	542	2	Q6GZAT_7BARHE	Q6gzat bartonella	373	38	61.3	56	2	Q7UUH4_RHOBA	Q7uuh4 rhodopirell
301	38	61.3	553	2	Q4XSW6_PLACH	Q4xsw6 plasmodium	374	38	61.3	84	2	Q86W19_HUMAN	Q86w19 homo sapien
302	38	61.3	553	2	Q726F6_DESVH	Q726f6 desulfovibr	375	38	61.3	86	2	Q58N03_9CAUD	Q58n03 cyanophaga
303	38	61.3	575	2	Q24662_DROSU	Q24662 drosophila	376	38	61.3	86	2	Q8EL67_OCEIH	Q8el67 oceanobacil
304	38	61.3	575	2	Q24663_DROSU	Q24663 drosophila	377	38	61.3	99	1	RPOZ_ENTFA	Q82286 enterococcu
305	38	61.3	586	2	Q6FN57_CANGA	Q6fn57 candida gla	378	38	61.3	101	2	Q4SYX2_TETNG	Q4syx2 tetraodon n
306	38	61.3	609	1	THI3_YEAST	Q07471 saccharomyc	379	38	61.3	102	2	Q16462_CABEL	Q16462 caenorhabdi
307	38	61.3	612	2	Q9HU20_PSEAE	Q9hu20 pseudomonas	380	38	61.3	102	2	Q9M2L5_7ARATH	Q9m2l5 arabidopsis
308	38	61.3	616	2	Q4Z1U2_PLABE	Q4z1u2 plasmodium	381	38	61.3	103	2	Q4SY73_TETNG	Q4sy73 tetraodon n
309	38	61.3	639	2	Q5L795_CHLAB	Q5l795 chlamyidophi	382	38	61.3	112	2	Q7V279_PROMP	Q7v279 prochloroco
310	38	61.3	664	2	Q8PSJ9_METWA	Q8psj9 methanosaarc	383	38	61.3	121	2	Q4HK66_CAMLA	Q4hk66 campylobact
311	38	61.3	667	2	Q7RDN9_PLAYO	Q7rdn9 plasmodium	384	38	61.3	130	2	Q8Y9W2_LISMO	Q8y9w2 listeria mo
312	38	61.3	669	2	Q9YFF9_AERPE	Q9yff9 aeropyrum p	385	38	61.3	135	2	Q6IVM1_BRARE	Q6ivm1 brachydanio
313	38	61.3	690	2	Q7Z5W8_HUMAN	Q7z5w8 homo sapien	386	38	61.3	137	1	URE2_STAFY	Q42874 staphylococ
314	38	61.3	704	2	Q5CTH2_CRYPV	Q5cth2 cryptospori	387	38	61.3	141	2	Q5SIW0_THET8	Q5siw0 thermus the
315	38	61.3	758	2	Q4RUV9_TETNG	Q4ruv9 tetraodon n	388	38	61.3	141	2	Q6YQ93_ONYPE	Q6yq93 onion yello
316	38	61.3	762	1	MUTS2_HELPJ	Q92ll4 helicobacte	389	38	61.3	152	2	Q7UJ89_THET8	Q7uj89 bacterioph
317	38	61.3	762	1	MUTS2_HELPJ	Q25338 helicobacte	390	38	61.3	163	2	Q4ZEN8_9VIRU	Q4zcn8 bacterioph
318	38	61.3	778	2	Q5RFF9_AERPE	Q5rff9 aeropyrum p	391	38	61.3	165	2	Q4ZCY5_9VIRU	Q4zcy5 bacterioph
319	38	61.3	779	2	Q5RFM4_PONPY	Q5rfm4 pongo pygma	392	38	61.3	168	2	Q87EX0_XYLFT	Q87ex0 xylella fas
320	38	61.3	795	2	Q5ZL11_CHICK	Q5zl11 gallus gall	393	38	61.3	168	2	Q9PGT2_XYLPF	Q9pgt2 xylella fas
321	38	61.3	818	2	Q7R911_PLAYO	Q7r911 plasmodium	394	38	61.3	173	1	PYRR_STRP3	Q8k7y5 streptococc
322	38	61.3	829	2	Q6CMM1_KLULA	Q6cmm1 kluyveromyc	395	38	61.3	173	1	PYRR_STRP6	Q5xc51 streptococc
323	38	61.3	833	2	Q57YQ5_9TRYP	Q57yq5 trypanosoma	396	38	61.3	173	1	PYRR_STRP8	P59013 streptococc

397 37 59.7 173 1 PYRR_STRPV
398 37 59.7 178 2 Q4X7P8_PLACH
399 37 59.7 182 2 Q6MU10_MYCMA
400 37 59.7 184 2 Q6KI177_MYCMA
401 37 59.7 186 2 Q4XUD1_PLACH
402 37 59.7 189 2 Q6Q8U3_9GAMM
403 37 59.7 193 2 Q4YJ85_PLABE
404 37 59.7 200 2 Q7QAP1_ANOGA
405 37 59.7 211 2 Q5GRQ5_WOLTR
406 37 59.7 216 2 Q4XRW0_PLACH
407 37 59.7 216 2 Q6DF11_XENTR
408 37 59.7 222 2 Q4Z1Q4_PLABE
409 37 59.7 227 2 Q55133_SYNY3
410 37 59.7 230 2 Q7AN21_NANEQ
411 37 59.7 237 2 Q63GU2_BACCZ
412 37 59.7 249 2 Q5U903_PIG
413 37 59.7 253 2 Q7S9T4_NEUCR
414 37 59.7 253 2 Q4HDZ5_CAMCO
415 37 59.7 254 1 RS3A_YEAST
416 37 59.7 254 1 RS3B_YEAST
417 37 59.7 255 2 Q6CWD0_KLULA
418 37 59.7 255 2 Q6FK63_CANGA
419 37 59.7 256 2 Q756Y3_ASHGO
420 37 59.7 259 2 Q5WJ11_BACSK
421 37 59.7 266 2 Q5CMC7_CRYPV
422 37 59.7 266 2 Q4YI63_PLABE
423 37 59.7 266 2 Q5CI75_CRYHO
424 37 59.7 267 2 Q5UF70_9PROT
425 37 59.7 269 2 Q8TZQ4_PYRFU
426 37 59.7 272 2 Q15800_PLAFA
427 37 59.7 272 2 Q81B27_PLAF7
428 37 59.7 275 2 Q8WNJ5_DICDI
429 37 59.7 281 2 Q74JP9_LACJO
430 37 59.7 283 2 Q6M0F1_METMP
431 37 59.7 283 2 Q7RDJ9_PLAYO
432 37 59.7 291 2 Q9SQ78_ARATH
433 37 59.7 294 2 Q94QJ1_ARATH
434 37 59.7 294 2 Q9SWZ8_ARATH
435 37 59.7 300 2 Q7RS26_PLAYO
436 37 59.7 301 2 Q9SMQ1_ARATH
437 37 59.7 310 2 Q9M4Q1_BRARP
438 37 59.7 311 1 ATHB6_ARATH
439 37 59.7 311 2 Q69G85_BRANA
440 37 59.7 319 2 Q7TR83_MOUSE
441 37 59.7 319 2 Q8VFL2_MOUSE
442 37 59.7 331 2 Q7QHS5_ANOGA
443 37 59.7 342 1 YM22_YEAST
444 37 59.7 348 2 Q9LFT2_ARATH
445 37 59.7 357 2 Q7UA79_SYNPX
446 37 59.7 364 2 Q16245_CAEEL
447 37 59.7 370 2 Q7PX02_ANOGA
448 37 59.7 371 2 Q97JW8_CLOAB
449 37 59.7 372 2 Q64WZ0_BACFR
450 37 59.7 373 2 Q9XYH2_TOXGO
451 37 59.7 381 2 Q8D2Z5_WIGBR
452 37 59.7 390 2 Q74WC8_NANEQ
453 37 59.7 397 1 Y636_METJA
454 37 59.7 407 2 Q5UTJ23_ORYSA
455 37 59.7 411 2 Q5OTJ7_ENTHI
456 37 59.7 414 2 Q8D4L0_VIBVU
457 37 59.7 414 2 Q7WG44_VIBVY
458 37 59.7 419 2 Q4XDT2_PLACH
459 37 59.7 427 2 Q4Z334_PLABE
460 37 59.7 427 2 Q8RGL1_FUSNN
461 37 59.7 431 2 Q9E2L7_CYTJO
462 37 59.7 435 2 Q6LZ75_METMP
463 37 59.7 441 2 Q4SVI9_TETNG
464 37 59.7 443 2 Q7SDS3_ASHGO
465 37 59.7 444 2 Q7P2F5_FUSNV
466 37 59.7 446 2 Q73QP4_TREDE
467 37 59.7 446 2 Q7N043_PHOLL
468 37 59.7 451 1 NUF2_YEAST
469 37 59.7 451 2 Q6BLX5_YEAST

Q9a0d0 streptococc
Q4x7p8 plasmodium
Q6mu10 mycoplasma
Q6ki177 mycoplasma
Q4xud1 plasmodium
Q6q8u3 uncultured
Q4y785 plasmodium
Q7gap1 anopheles g
Q5grq5 wolbachia s
Q4xrwo plasmodium
Q6df11 xenopus tro
Q4z1q4 plasmodium
Q55133 synechocyst
Q7an27 nanoarchaeu
Q63gu2 bacillus ce
Q5u903 sus scrofa
Q7s9t4 neurospora
Q4hdz5 campylobact
P33442 saccharomyc
P23248 saccharomyc
Q6cwo0 kluyveromyc
Q756y3 ashbya goss
Q5wa11 bacillus cl
Q5cmc7 cryptospori
Q4yt63 plasmodium
Q5ci75 cryptospori
Q5uf70 uncultured
Q8tzq4 pyrococcus
Q15800 plasmodium
Q81b27 plasmodium
Q8mnj5 dictyosteli
Q74jp9 lactobacilli
Q6m0f1 methanococc
Q7rdj9 plasmodium
Q9sq78 arabidopsis
Q94qj1 arabidopsis
Q9swz8 arabidopsis
Q7rs26 plasmodium
Q9smq1 arabidopsis
Q9m4q1 brassica ra
Q46668 arabidopsis
Q69g85 brassica na
Q7tr83 mus musculu
Q8vfl2 mus musculu
Q7qhs5 anopheles g
P40214 saccharomyc
Q9lft2 arabidopsis
Q7ua79 synechococc
Q16245 caenorhabdi
Q7px02 anopheles g
Q97jw8 clostridium
Q64wz0 bacteroides
Q9xyh2 toxoplasma
Q8d2z5 wigleswort
Q74mc8 nanoarchaeu
Q58053 methanococc
Q4jft23 oryza sativ
Q50tj7 entamoeba h
Q8d4l0 vibrio vuln
Q7wg44 vibrio vuln
Q4xdt2 plasmodium
Q4z334 plasmodium
Q8rgl1 fusobacteri
Q9ezl7 cytophaga j
Q6lzt75 metprococ
Q4svi9 tetraodon n
Q7sd33 ashbya goss
Q7p2f5 fusobacteri
Q73qp4 treponema d
Q7n043 photorhabdu
P33895 saccharomyc
Q6blx5 saccharomyc

RESULT 1

SEPR_HUMAN
ID SEPR_HUMAN STANDARD; PRT: 760 AA.
AC Q12884; O00199; Q86229; Q99998; Q9UD4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Seprase (SC 3.4.21.-) (Fibroblast activation protein alpha) (Integral
DE membrane serine protease) (170-kDa melanoma membrane-bound
DE Gelatinase).
GN Name:FAP;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC TISSUE=Fibroblast;
RX MEDLINE=94261645; PubMed=7911242;
RA Scanlan M.J., Raj B.K.W., Calvo B., Garin-Chesa P., Sanz-Moncasi M.P.,
Healey J.H., Old L.J., Rettig W.J.;
RT "Molecular cloning of fibroblast activation protein alpha, a member of
RT the serine protease family selectively expressed in stromal
RT fibroblasts of epithelial cancers."
RL Proc. Natl. Acad. Sci. U.S.A. 91:5657-5661(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT THR-354.
RC TISSUE=Melanoma;
RX MEDLINE=97388251; PubMed=9247085; DOI=10.1016/S0925-4439(97)00032-X;
RA Goldstein L.A., Gherai G., Pineiro-Sanchez M.L., Salamone M., Yeh Y.,
Flessate D., Chen W.-T.;
RT "Molecular cloning of seprase: a serine integral membrane protease
RT from human melanoma."
RL Biochim. Biophys. Acta 1361:11-19 (1997).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND PROTEIN SEQUENCE OF 220-229;

ALIGNMENTS

461-472 AND 511-518.
TISSUE=Melanoma;
MEDLINE=97218181; PubMed=9065413; DOI=10.1074/jbc.272.12.7595;
Pinsiro-Sanchez M.L., Goldstein L.A., Dodi J., Howard L., Yeh Y.,
Chen W.-T.;
RT Identification of the 170-kDa melanoma membrane-bound gelatinase
(seprase) as a serine integral membrane protease.";
J. Biol. Chem. 272:7595-7601(1997).
[4]
RN NUCLEOTIDE SEQUENCE (ISOFORM 2).
RP TISSUE=Melanoma;
RC MEDLINE=20112818; PubMed=10644713; DOI=10.1074/jbc.275.4.2554;
Goldstein L.A., Chen W.-T.;
RA "Identification of an alternatively spliced seprase mRNA that encodes
a novel intracellular isoform.";
J. Biol. Chem. 275:2554-2559(2000).
[5]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
RP THR-354.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M.J., Udwin T.B., Toshiyuki S., Carninci P., Prance C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
RN PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.
RX MEDLINE=94327249; PubMed=7519594;
Rettig W.J., Su S.L., Fortunato S.R., Scanlan M.J., Raj B.K.M.,
Garin-Chesa P., Healey J.H., Old L.J.;
RA "Fibroblast activation protein: purification, epitope mapping and
induction by growth factors.";
Int. J. Cancer 58:385-392(1994).
CC -!- FUNCTION: May have a role in tissue remodeling during development
and wound healing, and may contribute to invasiveness in malignant
cancers.
CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and
type IV collagen, but not native type I or type IV collagen. Does
not cleave laminin, fibronectin, fibrin or casein.
CC -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
inactive.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell
surface lamellipodia, invadopodia and on shed vesicles.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1; Synonyms=L1;
IsoId=Q12884-1; Sequence=Displayed;
Note=Major isoform;
Name=2; Synonyms=S, Truncated;
IsoId=Q12884-2; Sequence=VSP_005367;
CC -!- TISSUE SPECIFICITY: Fibroblast specific.
CC -!- INDUCTION: In fibroblasts at times and sites of tissue remodeling
during development, tissue repair, and carcinogenesis.
CC -!- PTM: N-glycosylated.
CC -!- PTM: The N-terminus may be blocked.
CC -!- SIMILARITY: Belongs to the peptidase S9B family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

CC EMBL; U09278; AB49652.1; -; mRNA.
CC EMBL; U76833; AAC51668.1; -; mRNA.
CC EMBL; AF007822; AAF21600.1; -; mRNA.
CC EMBL; BC026250; AAF26250.1; -; mRNA.
CC PDB; 1Z68; X-ray; A=39-757, B=-.
CC MEROPS; S09.007; -.
CC Ensembl; ENSG00000078098; Homo sapiens.
CC HGNC; HGNC:3590; FAP.
CC MIM; 600403; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0030027; C:lamellipodium; IDA.
CC GO; GO:0005886; C:plasma membrane; NAS.
CC GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.
CC GO; GO:0004222; F:metalloendopeptidase activity; TAS.
CC GO; GO:0042803; F:protein homodimerization activity; NAS.
CC InterPro; IPR002471; Pept_S9_AS.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_estrs.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase_S9; 1.
CC PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW 3D-structure; Alternative splicing; Direct protein sequencing;
KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
KW Signal-anchor; Transmembrane.
FT TOPO_DOM 1 4 Cytoplasmic (Potential).
FT TRANSMEM 5 25 Signal-anchor for type II membrane
protein (Potential).
FT TOPO_DOM 26 760 Extracellular (Potential).
FT ACT_SITE 624 624 Charge relay system (By similarity).
FT ACT_SITE 702 702 Charge relay system (By similarity).
FT ACT_SITE 734 734 Charge relay system (By similarity).
FT CARBOHYD 49 49 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 92 92 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 1 521 Missing (in isoform 2).
FT VARIAT 354 354 R -> T (in dbSNP:1126507).
FT CONFLICT 207 207 P -> A (in Ref. 2 and 5).
FT CONFLICT 229 229 K -> T (in Ref. 2 and 5).
SQ SEQUENCE 760 AA; 87821 MW; A0D34B4801BE07EA CRC64;
Query Match 100.0%; Score 62; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.4; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;
QY 1 ILEENKELENALK 13 PRT; 760 AA.
Db 487 ILEENKELENALK 499
RESULT 2
Q53TP5 HUMAN PRELIMINARY; PRT; 760 AA.
AC Q53TP5
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FAP.
GN Name=FAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Cotton M., Maupin R., Hawkins M., Harkins R.;
 RT "The sequence of Homo sapiens BAC clone RP11-576116."
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RN NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AC007750; RAY24205.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR00379; Ser esters.
 DR Pfam; PF00930; DPPIV_N; I.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolyase; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 760 AA; 87713 MW; 7FF81785A4F75142 CRC64;

 Query Match 100.0%; Score 62; DB 2; Length 760;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 ILEENKELENALK 13
 |||||
 DB 487 ILEENKELENALK 499

 RESULT 3
 Q8R492 RAT PRELIMINARY; PRT; 761 AA.
 AC Q8R492;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Fibroblast activation protein alpha subunit.
 GN Name=Fap;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Zuckersweldt K., Park J.E.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF493782; AAM11677.1; -; mRNA.
 DR HSP; P27487; 1PFQ.
 DR MEROPS; S09.007; -.
 DR Ensembl; ENSRNOG0000005679; Rattus norvegicus.
 DR RGD; 621253; Fap.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002471; Pept_S9_AS.

DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR00379; Ser esters.
 DR Pfam; PF00930; DPPIV_N; I.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Hydrolyase; Transmembrane.
 SQ SEQUENCE 761 AA; 87843 MW; EB214BB84BB14F2C CRC64;

 Query Match 93.5%; Score 58; DB 2; Length 761;
 Best Local Similarity 84.6%; Pred. No. 1.8;
 Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 QY 1 ILEENKELENALK 13
 :|||||
 DB 487 VLEENKELENALR 499

 RESULT 4
 SEPR_MOUSE STANDARD; PRT; 761 AA.
 ID SEPR_MOUSE
 AC P97321;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Sepase (EC 3.4.21.-) (Fibroblast activation protein alpha) (Integral membrane serine protease).
 DE Name=Fap;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2 AND 3).
 RC STRAIN=BALB/c; TISSUE=Embryo;
 RX MEDLINE=97284459; PubMed=9139873;
 RA Niedermeyer J., Scanlan M.J., Garin-Chesa P., Daiber C., Fiebig H.H.,
 RA Old L.J., Rettig W.J., Schnapp A.;
 RT "Mouse fibroblast activation protein: molecular cloning, alternative splicing and expression in the reactive stroma of epithelial cancers.";
 RT Int. J. Cancer 71:383-389(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RC TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May have a role in tissue remodeling during development and wound healing, and contribute to invasiveness in malignant cancers.
 CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and type IV collagen, but not native type I or type IV collagen. Does not cleave laminin, fibronectin, fibrin or casein.

```

CC -|- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is
CC inactive (By similarity).
CC -|- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P97321-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P97321-2; Sequence=VSP_005368;
CC Name=3;
CC IsoId=P97321-3; Sequence=VSP_005369;
CC -|- TISSUE SPECIFICITY: Detected in fibroblasts, in placenta, uterus,
CC embryos from day 7-19 and in new-born mice (Pl).
CC -|- SIMILARITY: Belongs to the peptidase S9B family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; Y10007; CAA71116.1; -; mRNA.
CC EMBL; BC019190; AAH19190.1; -; mRNA.
CC HSSP; P27487; IPFO.
CC MEROPS; S09.007; -.
CC Ensembl; ENSMUSG00000000392; Mus musculus.
CC MGI; MGI:109608; Fap.
CC GO; GO:0016021; C:integral to membrane; TAS.
CC InterPro; IPR002471; Pept_S9_AS.
CC InterPro; IPR001375; Peptidase_S9.
CC InterPro; IPR002469; Peptidase_S9B.
CC InterPro; IPR000379; Ser_estra.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase_S9; 1.
CC PROSITE; PS00708; PRO_ENDOPSP_SER; 1.
CC KW Alternative splicing; Glycoprotein; Hydrolase; Protease;
KW Serine protease; Signal-anchor; Transmembrane.
FT TOPO_DOM 1 4 Cytoplasmic (Potential).
FT TRANSMEM 5 25 Signal-anchor for type II membrane
FT protein (Potential).
FT TOPO_DOM 26 761 Extracellular (Potential).
FT ACT_SITE 624 624 Charge relay system (By similarity).
FT ACT_SITE 702 702 Charge relay system (By similarity).
FT ACT_SITE 734 734 Charge relay system (By similarity).
FT CARBOHYD 49 49 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 92 92 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 227 227 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
FT VARSPIC 31 63 Missing (in isoform 3).
FT VARSPIC 31 35 /FridaVSP_005369.
FT CONFLICT 737 737 Missing (in isoform 2).
FT CONFLICT 737 737 /FridaVSP_005368.
FT SEQUENCE 761 AA; 87945 MW; 9174C3AEDA213B25 CRC64;
Query Match 88.7%; Score 55; DB 1; Length 761;
Best Local Similarity 76.9%; Pred. No. 5.3;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB 487 VLEENKELENSLR 499
-----
RESULT 5
QSCDS3_CRYHO PRELIMINARY; PRT; 352 AA.
ID QSCDS3_CRYHO PRELIMINARY;
AC QSCDS3;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)

```

```

DE Hypothetical protein.
GN ORFNames=chro.70572;
OS Cryptosporidium hominis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=237895;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TU502;
RA Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G.,
RA Puiu D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H.,
RA Bankier A.T., Peterson D.L., Abrahamson M.S., Kapur V., Tzipori S.,
RA Buck G.A.;
RT "The genome of Cryptosporidium hominis.";
RL Nature 431:1107-1112(2004).
DR EMBL; AAEL01000939; BAU34754.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 352 AA; 41011 MW; 83BE269AD8026310 CRC64;
Query Match 80.6%; Score 50; DB 2; Length 352;
Best Local Similarity 90.9%; Pred. No. 15;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LEENKELENAL 12
DB 183 LEENKELENAL 193
-----
RESULT 6
QSCXS4_CRYPV PRELIMINARY; PRT; 1624 AA.
ID QSCXS4_CRYPV PRELIMINARY;
AC QSCXS4;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Hypothetical low complexity protein with coiled coil regions.
GN ORFNames=cgd7_5130;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahamante J.E., Zhu G.,
RA Lantto C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
RT parvum.";
RL Science 304:441-445(2004).
CC -|- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAEE01000001; BAK90470.1; -; Genomic_DNA.
DR InterPro; IPR000074; ApoA1_A4_E.
DR InterPro; IPR008456; Collagen_bind.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR000888; dTDP_Sugar_isom.
DR InterPro; IPR011992; EF-Hand_type.
DR InterPro; IPR000259; Fimbrial.
DR InterPro; IPR010486; HdeA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001783; Lum_binding.
DR InterPro; IPR009032; Vpu_cyt.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR PROSITE; PS00693; LUM_BINDING; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1624 AA; 190738 MW; F249813409CFE8E3 CRC64;
Query Match 90.6%; Score 50; DB 2; Length 1624;
Best Local Similarity 90.9%; Pred. No. 74;

```

```

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2  LEENKELENAL 12
      |||||:|||||
Db      1455 LEENKELENAL 1465

```

[illegible]

Query Match 79.0%; Score 49; DB 2; Length 653;
Best Local Similarity 69.2%; Pred. No. 41;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 ILEENKELENALK 13
        :||||:||||:
Db      83 VLEENQSLLENALR 95
```

```

RESULT 8
Q5XYQ4_BORGA
ID Q5XYQ4_BORGA PRELIMINARY; PRT; 147 AA.
AC Q5XYQ4;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BG227;
OS Borrelia garinii pB1.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=290434;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pb1;
RA Glockner G., Romualdi A., Pradella S.,
RA Schultze-Spechtel U., Schilhabel M., Wilske B., Suhnel J., Platzer M.;
RT "Comparative analysis of the Borrelia garinii genome.";
RL Nucleic Acids Res. 32:6038-6046(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pb1;
RA Gloeckner G., Schilhabel M., Lehmann R., Platzer M.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722929; AAU86078.1; -; Genomic DNA.
DR GO; GO:0006313; P:DNA transposition; IEA.
KW Hypothetical protein.
SO SEQUENCE 147 AA; 17320 MW; 2FB3455E2732018E CRC64;

```

Query Match	75.8%	Score 47;	DB 2;	Length 147;
Best Local Similarity	69.2%;	Pred. No. 18;		
Matches	9:	Conservative	3:	Mismatches
				1: Indels
				0: Gaps

Qy 1 ILEENKELENALK 13
: : | | | | : | | | |
pb 27 VIEENKILKNALK 39

RESULT 9

```

NIC1_SCHPO
ID NIC1_SCHPO STANDARD; PRT; 405 AA.
AC O74869; O74912;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE High-affinity nickel transporter protein nicl.
DE Name:nicl; ORFNames=SPCC1884.02, SPCC757.01;
GN Schizosaccharomyces pombe (Fission yeast).
OS Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA Sgouras J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K.D., Jones M., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney K., Moutle S., Mungall K.L., Murphy L.D., Niblett D., Odeh C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J.R., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rigier M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Duesterhoeft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Bozzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst A.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RT Nature 415:871-880(2002).
RN [2]
RP FUNCTION.
RX MEDLINE=20309746; PubMed=10748059; DOI=10.1074/jbc.M001978200;
RA Eitinger T., Degen O., Bohnke U., Muller M.;
RT "Nic1p, a relative of bacterial transition metal permeases in
RT Schizosaccharomyces pombe, provides nickel ion for urease
RT biosynthesis.";
RT J. Biol. Chem. 275:18029-18033(2000).
CC -1- FUNCTION: High-affinity nickel transporter responsible for nickel
CC uptake. Required for urease biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: Belongs to the NiCoF transporter (TC 2.A.52) family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL0311787; CA221158.1; -; Genomic DNA.
CC DR EMBL; AL031825; CA221226.1; -; Genomic DNA.
CC DR PIR; T41194; T41194.
CC DR GeneDB Spombe; SPCC1884.02; -.

```

```
DR InterPro: IPR011541; NcoI_transpept.
DR Pfam: PF03824; NcoI; 1.
KW Complete proteome; Nickel; Transmembrane; Transport.
FT TRANSMEM 35 55 Potential.
FT TRANSMEM 103 123 Potential.
FT TRANSMEM 138 158 Potential.
FT TRANSMEM 207 227 Potential.
FT TRANSMEM 235 255 Potential.
FT TRANSMEM 285 305 Potential.
FT TRANSMEM 330 350 Potential.
SQ SEQUENCE 405 AA; 45295 MW; 44CCB1BE7D13EFC8 CRC64;

Query Match 75.8%; Score 47; DB 1; Length 405;
Best Local Similarity 69.2%; Pred. No. 52;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 371 VLEKNKELENVSK 383
:|||||:

RESULT 10
Q97WW8 SULSO
ID Q97WW8_SULSO PRELIMINARY; PRT; 162 AA.
AC Q97WW8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SSO1993;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;

RN NUCLEOTIDE SEQUENCE.
RP Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Bertman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAJJ01007140; CAH84894.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 219 AA; 25862 MW; F44947037B46CAA6 CRC64;

Query Match 72.6%; Score 45; DB 2; Length 219;
Best Local Similarity 61.5%; Pred. No. 57;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 94 LFKENKELENKIK 106
:|||||:

RESULT 12
Q7WVB4_9STAP
ID Q7WVB4_9STAP PRELIMINARY; PRT; 176 AA.
AC Q7WVB4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative recombinase/mobilization protein.
OS Staphylococcus sciuri.
OG Plasmid pSCF51.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1296;
RN [1]_TaxID=1296;
RP NUCLEOTIDE SEQUENCE.
RA Kehrenberg C., Ojo K.K., Schwarz S.;
RT "Nucleotide sequence and organization of the multiresistance plasmid
RT pSCF51 from Staphylococcus sciuri.";
RL J. Antimicrob. Chemother. 54:936-939(2004).
DR EMBL; AJ579365; CAE18150.1; -; Genomic_DNA.
KW Plasmid.
SQ SEQUENCE 176 AA; 20896 MW; 6D1D22E205B8B891 CRC64;

Query Match 71.0%; Score 44; DB 2; Length 176;
Best Local Similarity 61.5%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 61 VYQENKELENK 73
:|||||:

RESULT 13
Q5Z181_CHICK
ID Q5Z181_CHICK PRELIMINARY; PRT; 759 AA.
AC Q5Z181;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_29g21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kuncic S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "full-length cDNAs from chicken bursal lymphocytes to facilitate
RT genefunction analysis.";
RL Genome Biol. 6:R6-R6(2005).
DR ENBL; AJ720903; CAG32562.1; -; mRNA.
DR CG; GO:0016020; C-membrane; IEA.
DR CG; GO:0004177; F:aminopeptidase activity; IEA.
DR CG; GO:0004374; F:diptidyl-peptidase IV activity; IEA.
DR CG; GO:0016787; F:hydrolase activity; IEA.
DR CG; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR CG; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF009930; DPPIV_N; I.
DR Pfam; PF003326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO-ENDOPEP SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 759 AA; 86343 MW; B9A0C38CCB8644AC CRC64;

Query Match 71.0%; Score 44; DB 2; Length 759;
Best Local Similarity 75.0%; Pred. No.3e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
||| |||||
Db 488 LENNTELESLK 499

RESULT 14
Q54BP1_DICDI PRELIMINARY; PRT; 2110 AA.
AC Q54BP1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80191988;
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Plicher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nle X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsegged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaubsky G., Schleicher M., Weinsrock G., Rosenthal T., Tanaka Y.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Nogel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum";
RL Nature 0:0-0(2005).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBS whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AF4101000293; EAL60584.1; -; Genomic DNA.
DR


```
RESULT 19
HDAL YEAST
ID HDAL YEAST STANDARD; PRT; 706 AA.
AC F53973;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Histone deacetylase HDAL
GN Name=HDAL; OrderedLocustNames=YNL021W; ORFNames=N2819;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]_TaxID=4932;
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c;
RX MEDLINE=97313269; PubMed=9169873;
RA Philippen P., Kleine K., Poehmann R., Duesterhoeft A., Hamberg K.,
RA Hegmann J.H., Obermaier B., Urrestazu L.A., Aert R., Albermann K.,
RA Altmann R., Andre B., Baladron V., Ballesta J.P.G., Becam A.-M.,
RA Beinbauer J.D., Boskovic J., Buirago M.J., Buserreau F., Coster F.,
RA Crouzet M., D'Angelo M., Dal Pero F., De Antoni A., del Rey F.,
RA Doignon F., Domdey H., Dubois E., Fiedler T.A., Fleig U., Floeth M.,
RA Fritz C., Gaillardin C., Garcia-Cantalejo J.M., Glansdorff N.,
RA Goffeau A., Guellder U., Herbert C.J., Heumann K., Heuss-Neitzel D.,
RA Hilbert H., Hinni K., Iraqui Housseini J., Jacquet M., Jimenez A.,
RA Joniaux J.-L., Karpfinger-Hartl L., Lanfranchi G., Lepingle A.,
RA Levesque H., Lyck R., Maftah M., Malliet L., Maurer C.T.C.,
RA Messenguy F., Mewes H.-W., Moestl D., Naer F., Nicaud J.-M.,
RA Niedenthal R.K., Pandolfo D., Pierard A., Piravandi E., Planta R.J.,
RA Pohl T.M., Purnelle B., Rebischung C., Remacha M.A., Revuelta J.L.,
RA Rinke M., Saiz J.E., Sartorillo F., Scherens B., Sen-Gupta M.,
RA Soler-Mira A., Urbane J.H.M., Valle G., Van Dyck L., Verhaesselt P.,
RA Vierendeels F., Viessers S., Voet M., Volckaert G., Wach A.,
RA Wambutt R., Wedler H., Zollner A., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIV
and its evolutionary implications.";
RL Nature 387:93-98(1997).
RN [2]
CHARACTERIZATION.
RX MEDLINE=97121415; PubMed=8962081; DOI=10.1073/pnas.93.25.14503;
RA Rundlett S.E., Carmen A.A., Kobayashi R., Bavykin S., Turner B.M.,
RA Grunstein M.;
RT "HDAL and RPD3 are members of distinct yeast histone deacetylase
complexes that regulate silencing and transcription.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:14503-14508(1996).
CC -1- FUNCTION: Responsible for the deacetylation of lysine residues on
the N-terminal part of the core histones (H2A, H2B, H3 and H4).
CC Histone deacetylation gives a tag for epigenetic repression and
plays an important role in transcriptional regulation, cell cycle
progression and developmental events. Histone deacetylases act via
the formation of large multiprotein complexes.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the histone deacetylase family. Type 2
subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
CC EMBL; Z71297; CAA95883.1; -; Genomic_DNA.
CC FIR; S62933; S62933.
CC IntAct; P53973; -.
CC GeneOnline; 143028; -.
CC TRANSFAC; T04600; -.
CC SGD; S00004966; HDAL.
CC Ensembl; YNL021W; Saccharomyces cerevisiae.
CC GO; GO:0004078; Histone deacetylase complex; IDA.
CC GO; GO:0004407; Histone deacetylase activity; IDA.
CC GO; GO:0006325; P:establishment and/or maintenance of chromatin.; TAS.
CC GO; GO:0001308; P:loss of chromatin silencing during replicat.; IMP.
-----
DR GO; GO:0006476; P:protein amino acid deacetylation; IDA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
DR InterPro; IPR000286; His_deacetylase.
DR PANTHER; PTHR10625; His_deacetylase; 1.
DR Pfam; PF00850; His_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Chromatin regulator; Complete proteome; Hydrolase; Nuclear protein;
KW Chromatin regulator; Transcription; Transcription regulation.
FT REGION 67 396 Histone deacetylase.
FT ACT_SITE 206 206 By similarity.
SQ SEQUENCE 706 AA; 80070 MW; 4E7069E6D03264D CRC64;
Query Match 69.4%; Score 43; DB 1; Length 706;
Best Local Similarity 81.8%; Pred. No. 4.1e+02;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 LEENKELENAL 12
DB 22 LEENKEEENSL 32
RESULT 20
Q9ZM29 HELPU
ID Q9ZM29 HELPU PRELIMINARY; PRT; 806 AA.
AC Q9ZM29;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative.
GN OrderedLocustNames=JHP0061;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
NUCLEOTIDE SEQUENCE.
RX MEDLINE=99120557; PubMed=9233682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Gullid B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001445; AAD05645.1; -; Genomic_DNA.
DR FIR; A71979; A71979.
DR InterPro; IPR002543; FtsK_SpoIIIE.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR PROSITE; PS0901; FTSK; 1.
KW Complete proteome.
SQ SEQUENCE 806 AA; 92770 MW; A08177FBA717590C CRC64;
Query Match 69.4%; Score 43; DB 2; Length 806;
Best Local Similarity 69.2%; Pred. No. 4.7e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
DB 155 ILTESKEIEEALK 167
RESULT 21
Q24906 HELPU
ID Q24906 HELPU PRELIMINARY; PRT; 831 AA.
AC Q24906;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved hypothetical ATP-binding protein.
GN OrderedLocustNames=HP0066;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
```

```

OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,
RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,
RA Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,
RA Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,
RA Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,
RA Smith H.O., Fraser C.M., Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547 (1997).
DR EMBL; AF000528; AAD07129.1; -; Genomic_DNA.
DR PIR; B64528; B64528.
DR TIGR; HP0066; -.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002543; FtsK_SpoIIIE.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS50901; FTSK; 1.
DR Complete proteome.
KW
SQ SEQUENCE 831 AA; 95646 MW; F674CBA4A82AF587 CRC64;

Query Match 69.4%; Score 43; DB 2; Length 831;
Best Local Similarity 69.2%; Pred. No. 4.8e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
DB 178 ILTSEKIEEALK 190
|||:|||||
|||:|||||

RESULT 22
OS 59254_PYRHO PRELIMINARY; PRT; 172 AA.
AC O59254;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein PH1595.
GN OrderedLocusNames=PH1595;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kusbida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76 (1998).
DR EMBL; BA000001; BAA30707.1; -; Genomic_DNA.
DR PIR; C71038; C71038.
DR InterPro; IPR00644; CBS.
DR Pfam; PF00571; CBS; 1.
DR SMART; SM00116; CBS; 2.
DR Complete proteome; Hypothetical protein.
KW
SQ SEQUENCE 172 AA; 19843 MW; 5F5FDBD05EEAE7CA CRC64;

Query Match 67.7%; Score 42; DB 2; Length 310;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 EENKELENAL 12
DB 32 EENKQFENAL 41
|||||
|||||

RESULT 24
OS Q66RN1_HEVR PRELIMINARY; PRT; 310 AA.
AC Q66RN1;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE MYB transcription factor.
GN Name=MYB;
OS Hevea brasiliensis (Para rubber tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae;
OC Micrandreae; Hevea.

Query Match 67.7%; Score 42; DB 2; Length 310;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 EENKELENAL 13
DB 30 ILEENSIINALK 42
|||||
|||||

RESULT 23
OS Q9QF4_HEVR PRELIMINARY; PRT; 310 AA.
AC Q9QF4;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Hevea brasiliensis (Para rubber tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae;
OC Micrandreae; Hevea.
OX NCBI_TaxID=3981;
RN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22490495; PubMed=12602890; DOI=10.1023/A:1020719420867;
RA Chen S., Peng S., Huang G., Wu K., Fu X., Chen Z.;
RT "Association of decreased expression of a Myb transcription factor
RT with the TPD (tapping panel dryness) syndrome in Hevea brasiliensis.";
RL Plant Mol. Biol. 51:51-58 (2003).
RN
RP NUCLEOTIDE SEQUENCE.
RA Chen S.C., Huang G.X., Feng S.Q.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AF239956; AAG44394.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001005; MYB_DNA_bd.
DR InterPro; IPR006447; SHAKYF_myb_bd.
DR Pfam; PF00249; MYB_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR TIGRFAMs; TIGR01557; myb_SHAKYF; 1.
DR PROSITE; PS50090; MYB_3_2.
DR Hypothetical protein; Nuclear protein; Repeat.
KW
SQ SEQUENCE 310 AA; 35059 MW; AE14624827D57FC6 CRC64;

```

```
OX NCBI_TaxID=3981;
RN [1] :|||||
RP NUCLEOTIDE SEQUENCE.
RC Venkatchalam P., Raghothama K.G., Thulaseedharan A.;
RT "Cloning and nucleotide sequence analysis of MYB transcription factor
RT associated with TPD syndrome in rubber tree (Hevea brasiliensis).";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AY112938; AA006309.1; -; Genomic DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001005; Myb DNA bd.
DR InterPro; IPR006447; SHAKRYF_myb bd.
DR Pfam; PF00249; Myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR TIGRFAMs; TIGR01557; myb SHAKRYF; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW Nuclear protein; Repeat.
SQ SEQUENCE 310 AA; 35058 MW; 87F7A24821D87B74 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 310;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 ENKELENAL 12
DB 32 EENKQFENAL 41

RESULT 25
Q8FR9 FUSNN PRELIMINARY; PRT; 342 AA.
AC Q8FR9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spermidine/putrescine-binding protein.
GN OrderedLocustNames=FN0618;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1] :|||||
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RX DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapacral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AB009951; AAL94814.1; -; Genomic DNA.
DR HSP; P21861; 1POT.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006910; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR InterPro; IPR001188; Sperm_putr_bd.
DR Pfam; PF01547; SBP_bac_1; 1.
DR PRINTS; PR00909; SPERNENDNG.
KW Complete proteome.
SQ SEQUENCE 342 AA; 39653 MW; 3FDEB024574FDC5F CRC64;

Query Match 57.7%; Score 42; DB 2; Length 342;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ILEENKELENALK 13
DB 261 VLKDSKHLENAMK 273

RESULT 26
Q86X02 HUMAN PRELIMINARY; PRT; 459 AA.
AC Q86X02;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Paraneoplastic antigen.
GN Name=HUMPPA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1] :|||||
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2] :|||||
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047534; AAH47534.1; -; mRNA.
DR Ensembl; ENSG00000109089; Homo sapiens.
SQ SEQUENCE 459 AA; 52368 MW; 375796AF0D7E764 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 459;
Best Local Similarity 61.5%; Pred. No. 3.7e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
DB 32 LLERNKELEGLQ 44

RESULT 27
Q4XEG4 PLACH PRELIMINARY; PRT; 523 AA.
AC Q4XEG4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC301191.00.0;
OC Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
```

```

RN NUCLEOTIDE SEQUENCE.
RP Hall N., Kariis M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole-genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: CAJ01007003; CAH84704.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 523 AA; 61763 MW; 761E46033CBBBC6 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 523;
Best Local Similarity 69.2%; Pred. No. 4.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
||| ||| ||| |||
DB 510 ILSNDLKNMLK 522

RESULT 28
Q6PNN2 HUMAN
ID Q6PNN2_HUMAN PRELIMINARY; PRT; 534 AA.
AC Q6PNN2
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Prostate cancer antigen T21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue-Testis;
RA Miles A.K., Li G., Rees R.C., Seth R., Darlison M.G., Bonner P.L.R.,
RA McArdle S.E.B., Mian S., Ali S.A.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY590151; AAT01278.1; -; mRNA.
SQ SEQUENCE 534 AA; 62429 MW; 44730BF32F13DB87 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 534;
Best Local Similarity 53.8%; Pred. No. 4.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:||| ||| ||| |||
DB 70 LVEENKQLEEGMK 82

RESULT 29
Q15175 HUMAN
ID Q15175_HUMAN PRELIMINARY; PRT; 535 AA.
AC Q15175;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Paraneoplastic antigen (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RA Fathallah-Shaykh H.M., Finizio J., Ho A., Rosenblum M., Posner J.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: L02867; AAA91850.1; -; mRNA.
DR Ensembl; ENSG00000109089; Homo sapiens.
FT NON_TER 1
SQ SEQUENCE 535 AA; 58013 MW; 6C80A459B888A1DA CRC64;

Query Match 67.7%; Score 42; DB 2; Length 535;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:||| ||| ||| |||
DB 108 LLERNKELEGSQ 120

RESULT 30
Q9H6Q7 HUMAN
ID Q9H6Q7_HUMAN PRELIMINARY; PRT; 720 AA.
AC Q9H6Q7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ21979.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK025632; BAB15196.1; -; mRNA.
FT NON_TER 720
SQ SEQUENCE 720 AA; 84029 MW; A86586FEAA953D0B CRC64;

Query Match 67.7%; Score 42; DB 2; Length 720;
Best Local Similarity 53.8%; Pred. No. 6e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:||| ||| ||| |||
DB 275 LVEENKQLEEGMK 287

RESULT 31
SYA_RICPR
ID SYA_RICPR STANDARD; PRT; 877 AA.
AC Q9ZCA4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Alanine-tRNA synthetase (EC 6.1.1.7) (Alanine--trna ligase) (AlaRS).
GN Name=alas; OrderedLocusNames=RP856;
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia; typhus group.
OX NCBI_TaxID=782;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893; DOI=10.1038/24094;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140 (1998).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +

```

```
CC      diphosphate + L-alanyl-tRNA(Ala).
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -|- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC      family.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AJ235273; CAA15280.1; -; Genomic_DNA.
CC      PIR; H71647; H71647.
CC      HAMAP; MF_00036; -; 1.
CC      InterPro; IPR003156; Pesterase_DHHAL.
CC      InterPro; IPR002318; tRNA-synt_2c.
CC      InterPro; IPR012947; tRNA_SAD.
CC      Pfam; PF02272; DHHAL; 1.
CC      Pfam; PF01411; tRNA-synt_2c; 1.
CC      Pfam; PF07973; tRNA_SAD; 1.
CC      PRINTS; PR00980; TRNASYNTHALA.
CC      TIGRFAMs; TIGR00344; alas; 1.
CC      PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
CC      Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW      Nucleotide-binding; Protein biosynthesis
SQ      SEQUENCE 877 AA; 99414 MW; 2BAB740998D4379B CRC64;

Query Match 67.7%; Score 42; DB 1; Length 877;
Best Local Similarity 69.2%; Pred. No. 7.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
DB      738 ILERNKELEKELE 750
      ||| ||||| :
      ||| ||||| :

RESULT 32
SYA RICTY STANDARD; PRT; 877 AA.
AC      Q68VQ7;
DT      10-MAY-2005 (Rel. 47, Created)
DT      10-MAY-2005 (Rel. 47, Last sequence update)
DT      13-SEP-2005 (Rel. 48, Last annotation update)
DE      Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlARS).
GN      Name=alas; OrderedLocusNames=RT0845;
OS      Rickettsia typhi.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC      Rickettsiaceae; Rickettsiae; Rickettsia; typhus group.
OX      NCBI_TaxID=785;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=Wilmington;
RX      PubMed=15317790; DOI=10.1126/JB.186.17.5842-5855.2004;
RA      McLeod M.P., Qin X., Karpathy S.E., Gioia J., Highlander S.K.,
RA      Fox G.E., McNeill T.Z., Jiang H., Muzny D., Jacob L.S., Hawes A.C.,
RA      Sodergren E., Gill R., Hume J., Morgan M., Fan G., Amin A.G.,
RA      Gibbs R.A., Hong C., Yu X.-J., Walker D.H., Weinstein G.M.;
RT      "Complete genome sequence of Rickettsia typhi and comparison with
RT      sequences of other Rickettsiae.";
RL      J. Bacteriol. 186:5842-5855(2004).
CC      -|- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC      diphosphate + L-alanyl-tRNA(Ala).
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -|- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC      family.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AB017197; AU04299.1; -; Genomic_DNA.
```

```
DR      HAMAP; MF_00036; -; 1.
DR      InterPro; IPR003156; Pesterase_DHHAL.
DR      InterPro; IPR002318; tRNA-synt_2c.
DR      InterPro; IPR012947; tRNA_SAD.
DR      Pfam; PF02272; DHHAL; 1.
DR      Pfam; PF01411; tRNA-synt_2c; 1.
DR      Pfam; PF07973; tRNA_SAD; 1.
DR      PRINTS; PR00980; TRNASYNTHALA.
DR      TIGRFAMs; TIGR00344; alas; 1.
DR      PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
KW      Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW      Nucleotide-binding; Protein biosynthesis
SQ      SEQUENCE 877 AA; 99125 MW; CF51798B7FC0016E CRC64;

Query Match 67.7%; Score 42; DB 1; Length 877;
Best Local Similarity 69.2%; Pred. No. 7.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
DB      738 ILERNKELEKELE 750
      ||| ||||| :
      ||| ||||| :

RESULT 33
Q4UJTS_RICFE PRELIMINARY; PRT; 878 AA.
ID      Q4UJTS_RICFE PRELIMINARY; PRT; 878 AA.
AC      Q4UJTS_5;
DT      13-SEP-2005 (TrEMBLrel. 31, Created)
DT      13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Alanyl-tRNA synthetase (EC 6.1.1.7).
GN      Name=alas; OrderedLocusNames=RF_1353;
OS      Rickettsia felis (Rickettsia azadi).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC      Rickettsiaceae; Rickettsiae; Rickettsia; spotted fever group.
OX      NCBI_TaxID=42862;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=URRWXcal2;
RX      PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA      Ogata H., Renesto P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA      Paxinello H., Claverie J.-M., Raoult D.;
RT      "The genome sequence of Rickettsia felis identifies the first putative
RT      conjugative plasmid in an obligate intracellular parasite.";
RL      PLOS Biol. 3:E248-E248(2005).
DR      EMBL; CP000053; AAY62204.1; -; Genomic_DNA.
DR      InterPro; IPR002048; EF_hand_Ca_bd.
DR      InterPro; IPR003156; Pesterase_DHHAL.
DR      InterPro; IPR002318; tRNA-synt_2c.
DR      InterPro; IPR012947; tRNA_SAD.
DR      Pfam; PF02272; DHHAL; 1.
DR      Pfam; PF01411; tRNA-synt_2c; 1.
DR      Pfam; PF07973; tRNA_SAD; 1.
DR      PRINTS; PR00980; TRNASYNTHALA.
DR      TIGRFAMs; TIGR00344; alas; 1.
DR      PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
DR      PROSITE; PS00018; EF_HAND; UNKNOWN_1.
KW      Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
SQ      SEQUENCE 878 AA; 99621 MW; CA941B9A40D10B98 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 878;
Best Local Similarity 69.2%; Pred. No. 7.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
DB      738 ILERNKELEKELE 750
      ||| ||||| :
      ||| ||||| :

RESULT 34
TSPI_BOVIN STANDARD; PRT; 1170 AA.
ID      TSPI_BOVIN STANDARD; PRT; 1170 AA.
AC      Q28178; Q28179;
```

DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP-1, TSPI;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_Nucleotide SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Holstein; Tissue=Tooth;
RX MEDLINE=98173773; PubMed=9507054; DOI=10.1016/S0167-4838(97)00188-X;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "CDNA cloning of bovine thrombospondin 1 and its expression in
odontoblasts and predentin.";
RT Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-18 AND 710-1170.
RC Tissue=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
TGF-beta.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibronectin, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-11b/beta-3. May play a role in dentinogenesis
and/or maintenance of dentin and dental pulp.
CC -1- SUBUNIT: Homotrimer; disulfide-linked.
CC -1- TISSUE SPECIFICITY: Odontoblasts.
CC -1- SIMILARITY: Belongs to the thrombospondin family.
CC -1- SIMILARITY: Contains 3 EGF-like domains.
CC -1- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -1- SIMILARITY: Contains 7 TSP type-3 domains.
CC -1- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; AB005287; BAA21115.1; -; mRNA.
DR EMBL; X87618; CAA60950.1; -; mRNA.
DR EMBL; X87619; CAA60951.1; -; mRNA.
DR PIR; S55501; S55501.
DR HSP; P07996; ILSL.
DR SNR; Q28178; 831-884, 834-1169.
DR GlycoSuiteDB; Q28178; -.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP-1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSP C.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF000008; EGF_1.
DR Pfam; PF000900; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; VWC; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS50092; TSPI; 3.
DR PROSITE; PS01208; VWFC 1; 1.
DR PROSITE; PS50184; VWFC_2; 1.

KW Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT CHAIN 1 18 By similarity.
FT DOMAIN 19 1170 Thrombospondin-1.
FT DOMAIN 24 221 TSP N-terminal.
FT DOMAIN 316 373 VWFC.
FT DOMAIN 379 429 TSP type-1 1.
FT DOMAIN 435 490 TSP type-1 2.
FT DOMAIN 492 547 TSP type-1 3.
FT DOMAIN 549 587 EGF-like 1.
FT DOMAIN 588 645 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 646 690 EGF-like 3.
FT DOMAIN 723 758 TSP type-3 1.
FT DOMAIN 759 781 TSP type-3 2.
FT DOMAIN 782 817 TSP type-3 3.
FT DOMAIN 818 840 TSP type-3 4.
FT DOMAIN 841 878 TSP type-3 5.
FT DOMAIN 879 914 TSP type-3 6.
FT DOMAIN 915 950 TSP type-3 7.
FT DOMAIN 951 1170 TSP C-terminal.
FT MOTIF 19 232 Heparin-binding (Potential).
FT CARBOHYD 248 248 Cell attachment site (Potential).
FT CARBOHYD 360 360 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 708 708 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1085 1085 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 270 Interchain (Probable).
FT DISULFID 274 274 Interchain (Probable).
FT DISULFID 301 423 By similarity.
FT DISULFID 395 428 By similarity.
FT DISULFID 406 413 By similarity.
FT DISULFID 447 484 By similarity.
FT DISULFID 451 489 By similarity.
FT DISULFID 462 474 By similarity.
FT DISULFID 504 541 By similarity.
FT DISULFID 508 546 By similarity.
FT DISULFID 519 531 By similarity.
FT DISULFID 551 562 By similarity.
FT DISULFID 556 572 By similarity.
FT DISULFID 575 586 By similarity.
FT DISULFID 592 608 By similarity.
FT DISULFID 599 617 By similarity.
FT DISULFID 620 644 By similarity.
FT DISULFID 650 663 By similarity.
FT DISULFID 657 676 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 705 713 By similarity.
FT DISULFID 718 738 By similarity.
FT DISULFID 754 774 By similarity.
FT DISULFID 777 797 By similarity.
FT DISULFID 813 833 By similarity.
FT DISULFID 836 856 By similarity.
FT DISULFID 874 894 By similarity.
FT DISULFID 910 930 By similarity.
FT DISULFID 946 1167 By similarity.
FT CONFLICT 805 805 S -> G (in Ref. 2).
SQ SEQUENCE 1170 AA; 129534 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 67.7%; Score 42; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.9e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANEUR 313

RESULT 35
TSP1_HUMAN
ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
AC P07996; Q15667;
DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Thrombospondin-1 precursor.
GN Name=THBS1; Synonyms=TSP, TSP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=endothelial cell;
RC MEDLINE=87057617; PubMed=2430973; DOI=10.1083/jcb.103.5.1635;
RX Lawler J., Hynes R.O.;
RA Baugartel D.M., Rotwein P., Frazier W.A.;
RA Haugartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RN NUCLEOTIDE SEQUENCE OF 1-397.
RP MEDLINE=87157592; PubMed=3030396;
RX Kobayashi S., Eden-Mecutchan P., Framson P., Bornstein P.;
RA "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RN NUCLEOTIDE SEQUENCE OF 1-374.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RN NUCLEOTIDE SEQUENCE OF 1-166.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RN NUCLEOTIDE SEQUENCE OF 1028-1170.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RT "Expression of thrombospondin in chronic inflammation: neutrophils
RT from synovial fluids synthesize a novel 3.9 kb TSP mRNA.";
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RN CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=platelet;
RX MEDLINE=21125860; PubMed=11067851; DOI=10.1074/jbc.M008073200;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.";
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RN THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399; DOI=10.1021/bi026463u;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 1 TSP C-terminal (TSPC) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M25631; AAA36741.1; -; mRNA.
CC EMBL; X04665; CAA28370.1; -; mRNA.
CC EMBL; X14787; CAA32889.1; -; mRNA.
CC EMBL; M14326; AAA61237.1; ALT_SEQ; mRNA.
CC EMBL; J04835; AAA61178.1; -; Genomic_DNA.
CC EMBL; M99425; AAB59366.1; -; mRNA.
CC PIR; A26155; TSHUP1.
CC PDB; 1LSL; X-ray; A=434-546.
CC PDB; LUX6; X-ray; A=834-1170.
CC GlycosuitedB; P07996; -.
CC OGP; P07996; -.
CC Ensemble; ENSG00000137801; Homo sapiens.
CC HGNC; HGNC:11785; THBS1.
CC Reactome; P07996; -.
CC MIM; 188060; -.
CC GO; GO:0005576; C:extracellular region; NAS.
CC GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
CC GO; GO:0004871; F:signal transducer activity; TAS.
CC GO; GO:0007275; P:development; TAS.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR003129; Laminin_G_TSP_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR000885; TSP1.
CC InterPro; IPR003367; tsg_3.
CC InterPro; IPR008859; TSP_C.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00090; TSP_1; 3.
CC Pfam; PF02412; TSP_3; 12.
CC Pfam; PF05735; TSP_C; 1.
CC Pfam; PF00093; VWFC_1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF_3.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWFC_1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS01084; VWFC_2; 1.
KW 3D-structure; Calcium; Cell adhesion; EGF-like domain; Glycoprotein;
KW Heparin-binding; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 Thrombospondin-1.
FT DOMAIN 24 221 TSP N-terminal.
FT DOMAIN 316 373 VWFC.
FT DOMAIN 379 429 TSP type-1 1.
FT DOMAIN 435 490 TSP type-1 2.
FT DOMAIN 492 547 TSP type-1 3.
FT DOMAIN 549 587 EGF-like 1.
FT DOMAIN 588 645 EGF-like 2; calcium-binding (Potential).
FT DOMAIN 646 690 EGF-like 3.


```
FT DOMAIN 723 758 TSP type-3 1.
FT DOMAIN 759 781 TSP type-3 2.
FT DOMAIN 782 817 TSP type-3 3.
FT DOMAIN 818 840 TSP type-3 4.
FT DOMAIN 841 878 TSP type-3 5.
FT DOMAIN 879 914 TSP type-3 6.
FT DOMAIN 915 950 TSP type-3 7.
FT DOMAIN 951 1170 TSP C-terminal.
FT REGION 19 232 Heparin-binding (Potential).
FT MOTIF 926 928 Cell attachment site (Potential).
FT CARBOHYD 248 248 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 360 360 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 385 385 C-linked (Man).
FT CARBOHYD 394 394 /FTId-CAR 000205.
FT CARBOHYD 394 394 O-linked (Fuc...).
FT CARBOHYD 438 438 /FTId-CAR 000206.
FT CARBOHYD 441 441 /FTId-CAR 000207.
FT CARBOHYD 450 450 /FTId-CAR 000208.
FT CARBOHYD 498 498 O-linked (Fuc...).
FT CARBOHYD 507 507 C-linked (Man).
FT CARBOHYD 507 507 /FTId-CAR 000210.
FT CARBOHYD 507 507 O-linked (Fuc...).
FT CARBOHYD 507 507 /FTId-CAR 000211.
FT CARBOHYD 708 708 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT DISULFID 270 274 Interchain (Probable).
FT DISULFID 391 423 Interchain (Probable).
FT DISULFID 406 413 Interchain (Probable).
FT DISULFID 447 484 By similarity.
FT DISULFID 451 489 By similarity.
FT DISULFID 556 572 By similarity.
FT DISULFID 575 586 By similarity.
FT DISULFID 592 608 By similarity.
FT DISULFID 599 617 By similarity.
FT DISULFID 620 644 By similarity.
FT DISULFID 650 663 By similarity.
FT DISULFID 657 676 By similarity.
FT DISULFID 678 689 By similarity.
FT DISULFID 705 713 By similarity.
FT DISULFID 718 738 By similarity.
FT DISULFID 754 774 By similarity.

Query Match 67.7%; Score 42; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 9.9e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313

RESULT 36
Q59E99 HUMAN PRELIMINARY; PRT; 1225 AA.
AC Q59E99;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Thrombospondin 1 variant (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
```

```
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE-Aorta endothelial cell;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB20912; BAD93149.1; -; mRNA.
DR SMR; Q59E99; 886-939, 889-1225.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSP_C.
DR InterPro; IPR001007; WVF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; TSP_1; 3.
DR Pfam; PF02412; TSP_3; 12.
DR Pfam; PF05735; TSP_C; 1.
DR Pfam; PF00093; WVC_1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WVF_C; 1.
DR PROSITE; PS01184; WVF_2; 1.
FT NON_TER 1
SQ SEQUENCE 1225 AA; 134849 MW; 9888B16857157B12 CRC64;

Query Match 67.7%; Score 42; DB 2; Length 1225;
Best Local Similarity 61.5%; Pred. No. 1e+03;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 356 VTEENKELANELR 368

RESULT 37
Q66GS8 HUMAN PRELIMINARY; PRT; 2479 AA.
AC Q66GS8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Centrosome protein Cep290.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=14654843; DOI=10.1038/nature02166;
RA Andersen J.S., Wilkinson C.J., Mayor T., Mortensen P., Nigg E.A.,
RA Mann M.;
RT "Proteomic characterization of the human centrosome by protein
correlation profiling."
RL Nature 426:570-574 (2003).
```



```

RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Mann M.;
RA Andersen J.S., Wilkinson C.J., Mayor T., Mortensen P., Nigg E.A.,
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BK005587; DAA05591.1; -; mRNA.
DR InterPro; IPR003900; KID repeat.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF02524; KID; 3.
SQ SEQUENCE 2479 AA; 290328 MW; 8183C7B843C1F49C CRC64;

Query Match 67.7%; Score 42; DB 2; Length 2479;
Best Local Similarity 53.8%; Pred. No. 2.2e+03;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 640 LVEENKLEEGMK 552
:::||||:|

RESULT 38
Q8KZ79 BORBU
ID Q8KZ79 BORBU PRELIMINARY; PRT; 73 AA.
AC Q8KZ79.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BppC.
OS Name=bppC;
GN Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-9.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Gwinn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Gwinn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Wattley L.A.,
RA Artach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
RL EMBL; AS001581; AAM60782.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005355; F:glucose transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
KW Complete proteome; Plasmid.
SQ SEQUENCE 73 AA; 8636 MW; FFF28C8E1A9984BF CRC64;

Query Match 66.1%; Score 41; DB 2; Length 73;
Best Local Similarity 61.5%; Pred. No. 80;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 640 LVEENKLEEGMK 552
:::||||:|

RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Mann M.;
RA Andersen J.S., Wilkinson C.J., Mayor T., Mortensen P., Nigg E.A.,
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BK005587; DAA05591.1; -; mRNA.
DR InterPro; IPR003900; KID repeat.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF02524; KID; 3.
SQ SEQUENCE 2479 AA; 290328 MW; 8183C7B843C1F49C CRC64;

Query Match 67.7%; Score 42; DB 2; Length 2479;
Best Local Similarity 53.8%; Pred. No. 2.2e+03;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 640 LVEENKLEEGMK 552
:::||||:|

RESULT 39
Q735K3 BACCI
ID Q735K3 BACCI PRELIMINARY; PRT; 126 AA.
AC Q735K3.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tn554-related, transposase C.
GN OrderedLocusNames=BCE3149, BCEA0002;
OS Bacillus cereus (strain ATCC 10987).
OC Plasmid pBci10987.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=222523;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PLASMID=pBci10987;
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Raako D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angluoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pX01."
RL Nucleic Acids Res. 32:977-988(2004)
DR EMBL; AB017274; AAS42059.1; -; Genomic_DNA.
DR EMBL; AB017195; AAS44852.1; -; Genomic_DNA.
DR TIGR; BCE3149; -.
DR TIGR; BCEA0002; -.
KW Complete proteome.
SQ SEQUENCE 126 AA; 14788 MW; 792D3CA4138F9A8D CRC64;

Query Match 66.1%; Score 41; DB 2; Length 126;
Best Local Similarity 72.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
Db 106 EENKQLRNQK 116
||||:|

RESULT 40
Q5BZ17 SCHJA
ID Q5BZ17 SCHJA PRELIMINARY; PRT; 140 AA.
AC Q5BZ17.
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS Schistosoma japonicum (blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY811299; AAX27188.1; -; mRNA.
DR InterPro; IPR010980; Cyt c.b562.
DR InterPro; IPR012290; Fibrinogen G.N.
DR InterPro; IPR011991; Wing_hlx_DNA_Dd.
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 16657 MW; 9376F30E197CE702 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 140;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
Db 76 EENKQLKNAIE 86
||||:|
```

```
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002104; Phage_integrase.
DR Pfam; PF00589; Phage_integrase; 1.
KW Plasmid.
SQ SEQUENCE 258 AA; 30273 MW; 7B65A7771A1D42A3 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 258;
Best Local Similarity 61.5%; Pred. No. 3e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 27 VLNENKILNSLK 39

RESULT 43
Q44780 BORBU
ID Q44780 BORBU PRELIMINARY; PRT; 258 AA.
AC Q44780; Q7DFC7;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE ORF-6 (Bppc)
GN Name=bppc; OrderedLocustNames=BBL38;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid cp32-1, and plasmid cp32-8.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B31; PLASMID=cp32-1;
RX MEDLINE=97136624; PubMed=8982001;
RA Casjens S., van Vugt R., Tilly K., Rosa P.A., Stevenson B.;
RT "Homology throughout the multiple 32-kilobase circular plasmids
present in Lyme disease spirochetes."
RL J. Bacteriol. 179:217-227(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B31; PLASMID=cp32-1;
RX MEDLINE=96256604; PubMed=8655548;
RA Stevenson B., Tilly K., Rosa P.A.;
RT "A family of genes located on four separate 32-kilobase circular
plasmids in Borrelia burgdorferi B31."
RL J. Bacteriol. 178:3508-3516(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B31; PLASMID=cp32-1;
RA Stevenson B.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8;
RX MEDLINE=20138354; PubMed=10672174;
RA Casjens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Hickey E.K.,
RA Hickey E.K., Winn M.L., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
extrachromosomal DNAs in an infectious isolate of the Lyme disease
spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31; PLASMID=cp32-8;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Matthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horet K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
```

```
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.", 586 (1997).
RL Nature 390:580-586 (1997).
DR EMBL; U44912; AAC34924.1; -; Genomic_DNA.
DR EMBL; AE001580; AAF07623.1; -; Genomic_DNA.
DR TIGR; BBL38; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0008313; P:DNA transposition; IEA.
DR InterPro; IPR002104; Phage integrase.
DR Pfam; PF00589; Phage integrase; 1.
KW Complete proteome; Plasmid.
SQ
SEQUENCE 258 AA; 30267 MW; DBF069AEDC1B6E58 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 258;
Best Local Similarity 61.5%; Pred. No. 3e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 27 VLNENKILKNSLK 39

RESULT 44
Q9S0J2 BORBU
ID Q9S0J2 BORBU PRELIMINARY; PRT; 258 AA.
AC Q9S0J2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BppC.
GN Name=bppC; OrderedLocustNames=BBP37;
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-1.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia;
OC Borrelia burgdorferi group.
OX NCBI_TaxID=139;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=ATCC 35210 / B31;
RC MEDLINE=20138354; PubMed=10672174;
RX Castens S., Palmer N., van Vugt R., Huang W.M., Stevenson B., Rosa P.,
RA Lathigra R., Sutton G.G., Peterson J.D., Dodson R.J., Haft D.H.,
RA Hickey E.K., Winn M.H., White O., Fraser C.M.;
RT "A bacterial genome in flux: the twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 35:490-516 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685; DOI=10.1038/37551;
RA Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R.J., Hickey E.K.,
RA Winn M.L., Dougherty B.A., Tomb J.-F., Fleischmann R.D.,
RA Richardson D.L., Peterson J.D., Kerlavage A.R., Quackenbush J.,
RA Salzberg S.L., Hanson M., van Vugt R., Palmer N., Adams M.D.,
RA Gocayne J.D., Weidman J.F., Utterback T.R., Watthey L., McDonald L.A.,
RA Artiach P., Bowman C., Garland S.A., Fujii C., Cotton M.D., Horst K.,
RA Roberts K.M., Hatch B., Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RL Nature 390:580-586 (1997).
DR EMBL; AB001575; AAF07426.1; -; Genomic_DNA.
DR TIGR; BBP37; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002104; Phage integrase.
DR Pfam; PF00589; Phage integrase; 1.
KW Complete proteome; Plasmid.
SQ
SEQUENCE 258 AA; 30279 MW; 5300785231178FA4 CRC64;
```

```
Query Match 66.1%; Score 41; DB 2; Length 258;
Best Local Similarity 61.5%; Pred. No. 3e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 27 VLNENKILKNSLK 39

RESULT 45
O80435 ARATH
ID O80435 ARATH PRELIMINARY; PRT; 291 AA.
AC O80435;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MYB family transcription factor.
GN Name=At2g38090;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AC003028; AAC27179.1; -; Genomic_DNA.
DR PIR; T01241; T01241.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001005; MYB_DNA_bd.
DR Pfam; PF00249; MYB_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR TIGRfam; TIGR01557; myb_SHAQYF; 1.
DR PROSITE; PS00090; MYB_3; 2.
KW Nuclear protein; Repeat.
SQ
SEQUENCE 291 AA; 32451 MW; CSBF5FB42E47ADDC CRC64;

Query Match 66.1%; Score 41; DB 2; Length 291;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 BENKELENAL 12
Db 25 BENKKFENAL 34

RESULT 46
Q720T1 LISMF
ID Q720T1 LISMF PRELIMINARY; PRT; 294 AA.
AC Q720T1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Transcriptional regulator Pocr, putative.
GN OrderedLocustNames=LMOF2365.1157;
OS Listeria monocytogenes (serotype 4b / strain F3265).
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=265669;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RL PubMed=15115801; DOI=10.1093/nar/gkh562;
```

RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
RA Forberger H.A., Iran B., Katharopoulos S., Wonderling L.D., Uhlich G.A.,
RA Bayles D.O., Luchansky J.B., Fraser C.M.;
RA "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
RT borne pathogen *Listeria monocytogenes* reveal new insights into the
RT core genome components of this species."; ;
RL Nucleic Acids Res. 32:2386-2395(2004).
DR TIGR; AF017325; AAT03933.1; -; Genomic_DNA.
DR EMBL; AF017325; AAT03933.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH ARAC.
DR PRINTS; PRO0032; HTHARAC.
DR SMART; SM00342; HTH ARAC.
DR PROSITE; PS00041; HTH ARAC FAMILY 1; 1.
DR PROSITE; PS01124; HTH ARAC FAMILY 2; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 294 AA; 33836 MW; 491415EB91961469 CRC64;
Query Match 66.1%; Score 41; DB 2; Length 294;
Best Local Similarity 72.7%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 3 BENKELENALK 13
DB 188 BENKEIRKALK 198
RESULT 47
O8Y7W9 LISMO PRELIMINARY; PRT; 294 AA.
AC O8Y7W9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Regulatory protein similar to *Salmonella typhimurium* Pocr
DE Protein.
GN OrderedLocusNames=lm01150;
OS *Listeria monocytogenes*.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaser P., Frangeul L., Buchrieser C., Ruaniok C., Amend A.,
RA Baquero F., Berche P., Blocher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madsen E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of *Listeria* species."; ;
RL Science 294:849-852(2001).
DR EMBL; AL591978; CAC99228.1; -; Genomic_DNA.
DR FRP; AF1218; AF1218.
DR HSSP; P27246; IBL0.
DR ListList; LM01150; -;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH ARAC.
DR PRINTS; PRO0032; HTHARAC.
DR SMART; SM00342; HTH ARAC.
DR PROSITE; PS00041; HTH ARAC FAMILY 1; 1.
DR PROSITE; PS01124; HTH ARAC FAMILY 2; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 294 AA; 33888 MW; 8ACF63B347B26349 CRC64;
Query Match 66.1%; Score 41; DB 2; Length 294;
Best Local Similarity 72.7%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 3 BENKELENALK 13
DB 188 BENKEIRKALK 198
RESULT 48
O8GXN7 ARATH PRELIMINARY; PRT; 298 AA.
AC O8GXN7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative MYB family transcription factor (MYB transcription
DE factor).
GN ORFNames=At2g38090;
OS *Arabidopsis thaliana* (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Qu L., Gu H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AK118135; BAC42760.1; -; mRNA.
DR EMBL; EF008122; RAP04107.1; -; mRNA.
DR EMBL; AY159529; AAS0999.1; -; mRNA.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001005; Myb DNA bd.
DR Pfam; PF00249; Myb DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR TIGRFAMs; TIGR01557; myb_SHAQKYF; 1.
DR PROSITE; PS00090; MYB 3; 2.
KW Nuclear protein; Repeat.
SQ SEQUENCE 298 AA; 33251 MW; 812F4E21D6D8B440 CRC64;
Query Match 66.1%; Score 41; DB 2; Length 298;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

[illegible]

DR PROSITE; PS00600; AA TRANSFER CLASS 3; 1.
KW Complete proteome; Pyridoxal phosphate.
SQ SEQUENCE 440 AA; 45909 MW; DA6FAF199BE29036 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 440;
Best Local Similarity 61.5%; Pred. No. 5.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 333 LLAEGKRIENALK 345

RESULT 52

Q8U3F5_PYRFU PRELIMINARY; PRT; 459 AA.

AC Q8U3F5_PYRFU PRELIMINARY; PRT; 459 AA.

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Putative glutamate aminotransferase.

GN OrderedLocusNames=PF0513;

OS Pyrococcus furiosus.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=2261;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;

RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

RL "The complete sequence of the Pyrococcus furiosus genome.";

RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent

CC aminotransferase family.

DR EMBL; AB010174; AAL80637.1; -, Genomic_DNA.

DR HSSP; P12995; 1QJ3.

DR GO; GO:0030170; F:pyridoxal phosphate binding; IEA.

DR GO; GO:0008483; F:transaminase activity; IEA.

DR InterPro; IPR005814; F:transaminase_3.

DR InterPro; IPR001760; Opsin.

DR PANTHER; PTHR11986; Aminotrans_3; 1.

DR Pfam; PF00202; Aminotran_3; 1.

DR PROSITE; PS00600; AA TRANSFER CLASS 3; 1.

DR PROSITE; PS00238; OPSIN; UNKNOWN 1.

KW Complete proteome; Pyridoxal phosphate.

SQ SEQUENCE 459 AA; 50554 MW; 4B7BD1DAE13D7DA0 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 459;

Best Local Similarity 69.2%; Pred. No. 5.4e+02;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 336 IIEENLLENALK 348

RESULT 53

Q4V6Y2_PLACH

ID Q4V6Y2_PLACH PRELIMINARY; PRT; 496 AA.

AC Q4V6Y2;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein (fragment).

GN ORFNames=PC000508.00.0;

OS Plasmodium chabaudi.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5825;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,

RA Berriaman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,

RA James K., Rutherford K., Harris B., Harris D., Churcher C.,

RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos P.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86 (2005).

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; CAAJ01000540; CAH75078.1; -, Genomic_DNA.

KW Hypothetical protein.

FT NON_TER 496

SQ SEQUENCE 496 AA; 57781 MW; 8E64CDCB26EC6C54 CRC64;

Query Match 66.1%; Score 41; DB 2; Length 496;

Best Local Similarity 81.8%; Pred. No. 5.9e+02;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13

DB 270 EENKENENDLK 280

RESULT 54

Q9X2C2_THEMEA

ID Q9X2C2_THEMEA PRELIMINARY; PRT; 546 AA.

AC Q9X2C2;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=TM1802;

OS Thermotoga maritima.

OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.

OX NCBI_TaxID=2336;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;

RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

RA McDonald L.A., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,

RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RT "Evidence for lateral gene transfer between Archaea and Bacteria from

RT genome sequence of Thermotoga maritima.";

RL Nature 399:323-329 (1999).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

DR EMBL; AB001817; AAD36865.1; -, Genomic_DNA.

DR PIR; G72210; G72210.

DR TIGR; TM1802; -.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 546 AA; 62704 MW; 02D0A08B66046F1B CRC64;

Query Match 66.1%; Score 41; DB 2; Length 546;

Best Local Similarity 61.5%; Pred. No. 6.5e+02;

Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 317 ILEESKKIESQLK 329

RESULT 55

O62319_CAEEL

ID O62319_CAEEL PRELIMINARY; PRT; 606 AA.

AC O62319;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein M04C3.2.

ORFNames=M04C3.2;

```

OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sucang R., Barriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Pey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Linday R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saio T., Buchrieser C.,
RA Wandrop A., Feider M., Thangaveilu M., Johnson D., Knights A.,
RA Loulaeeg H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinovitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kasp A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAFI01000269; EAL61150.1; -; Genomic_DNA.
DR Hypothetical protein.
KW
SQ SEQUENCE 784 AA; 91010 MW; A96EE5BC0FF445FA CRC64;

Query Match 66.1%; Score 41; DB 2; Length 784;
Best Local Similarity 61.5%; Pred. No. 9.5e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENAUX 13
DB 355 LLEENEELKNTNK 367
:||||:|:|:|

RESULT 58
Q9VIK6 DROME PRELIMINARY; PRT; 790 AA.
ID AC Q9VIK6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG1962-PA, isoform A (CG1962-pb, isoform b) (Ld41224p).
GN ORFNAMES=CG1962;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blajez R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Besen P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```


Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Flosier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hosin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milehina N.V., Mohanty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler J.F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RN Science 287:2185-2195(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celisner S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence.";
 RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celisner S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RN Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RN Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley *Drosophila* Genome Project;
 RA Celisner S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence.";
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Berkeley;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guatin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003667; AAF33912.1; -; Genomic_DNA.
 DR EMBL; AY051930; AAK39354.1; -; mRNA.
 DR IntAct; Q9VIR6; -;
 DR Ensembl; CG1962; Drosophila melanogaster.
 DR FlyBase; FBgn0032876; CG1962.
 SQ SEQUENCE 790 AA; 89904 MW; D58BADCC759607C2 CRC64;
 Query Match 66.1%; Score 41; DB 2; Length 790;
 Best Local Similarity 61.5%; Pred. NO. 9.5e-02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 DB 32 LLERNKELETTFIK 44

RESULT 59
 SYA_HELPJ STANDARD; PRT; 847 AA.
 ID SYA_HELPJ
 AC Q9ZJY5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Alanyl-tRNA synthetase [EC 6.1.1.7] (Alanine--tRNA ligase) (AlaRS).
 GN Name=alaS; OrderedLocNames=JHPL162;
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RN Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
 CC diphosphate + L-alanyl-tRNA(Ala).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
 CC family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AE001543; AAD06738.1; -; Genomic_DNA.
 DR F1R; F71842; F71842.
 DR HAMAP; MF_00036; -; 1.
 DR InterPro; IPR003156; Pesterase DHHA1.
 DR InterPro; IPR002318; tRNA-synt_2c.
 DR InterPro; IPR012947; tRNA_SAD.
 DR Pfam; PF02272; DHHA1; 1.
 DR Pfam; PF01411; tRNA-synt_2c; 1.
 DR Pfam; PF07973; tRNA_SAD; 1.
 DR PRINTS; PR00980; TRNASYNTHALA.
 DR TIGRfam; TIGR00344; alas; 1.
 DR PROSITE; PS50860; AA_TRNA_LIGASE II ALA; 1.
 KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Nucleotide-binding; Protein biosynthesis.
 SQ SEQUENCE 847 AA; 94978 MW; 714B10CALF92D8F1 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 847;
Best Local Similarity 88.9%; Pred. No. 1e+03;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EENKELENA 11
|||||:|
DB 699 EENKELKNA 707

Search completed: March 11, 2006, 12:25:11
Job time : 145.263 secs

RESULT 60
SYA_HELPY
ID SYA_HELPY STANDARD; PRT; 847 AA.
AC P56452;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Alanyl-tRNA synthetase [EC 6.1.1.7] (Alanine--tRNA ligase) (ALARS).
GN Name=alas; OrderedLocNames=HP1241;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,
RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,
RA Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,
RA Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,
RA Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,
RA Smith H.O., Fraser C.M., Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC -----
CC EMBL; AE000629; AAD08285.1; -; Genomic_DNA.
CC PIR; A64675; A64675.
CC TIGR; HP1241; -.
CC HAMAP; MF_00036; -; 1.
CC InterPro; IPR003156; Pesterase_DHA1.
CC InterPro; IPR002318; tRNA-synt_2c.
CC InterPro; IPR012947; tRNA_SAD.
CC Pfam; PF02272; DHA1; 1.
CC Pfam; PF01411; tRNA-synt_2c; 1.
CC Pfam; PF07973; tRNA_SAD; 1.
CC PRINTS; PR00980; TRNASYNTHALA.
CC TIGRFAMs; TIGR00344; alas; 1.
CC PROSITE; PS50860; AA tRNA LIGASE II ALA; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Nucleotide-binding; Protein biosynthesis.
SQ SEQUENCE 847 AA; 94700 MW; 56464EDB61FC8357 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 847;
Best Local Similarity 88.9%; Pred. No. 1e+03;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EENKELENA 11

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:25:33 ; Search time 29.0789 Seconds
(without alignments)
36.961 Million cell updates/sec

Title: US-10-774-242A-6
Perfect score: 62
Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/ECTUS_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	547	2	US-09-949-016-8330
2	62	100.0	760	1	US-08-230-491A-2
3	62	100.0	760	1	US-08-619-280A-2
4	62	100.0	760	1	US-08-940-391-2
5	62	100.0	760	2	US-09-265-606-2
6	55	88.7	761	2	US-09-518-550-26
7	43	69.4	346	2	US-10-012-819-82
8	43	69.4	706	2	US-09-538-092-649
9	42	67.7	825	1	US-09-949-002-482
10	42	67.7	1170	1	US-08-313-288B-20
11	42	67.7	1170	2	US-09-657-472-2
12	42	67.7	1170	2	US-09-949-002-350
13	41	66.1	317	2	US-09-270-767-42366
14	40	64.5	62	2	US-09-180-167A-7
15	40	64.5	62	2	US-09-033-524B-7
16	40	64.5	218	2	US-08-985-526-1
17	40	64.5	245	2	US-09-180-167A-33
18	40	64.5	245	2	US-09-033-524B-33
19	40	64.5	340	1	US-08-097-831-2
20	40	64.5	391	2	US-09-583-110-2803
21	40	64.5	403	2	US-09-107-433-3191
22	40	64.5	441	2	US-08-985-526-3
23	40	64.5	959	2	US-09-914-259-67
24	40	64.5	961	2	US-09-914-259-66
25	40	64.5	2285	2	US-09-308-375-2
26	40	64.5	2285	2	US-09-932-183A-2
27	39	62.9	411	2	US-09-253-701-1
					Sequence 8330, Ap
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 26, Appl
					Sequence 82, Appl
					Sequence 649, App
					Sequence 482, App
					Sequence 20, Appl
					Sequence 2, Appli
					Sequence 350, App
					Sequence 42366, A
					Sequence 7, Appli
					Sequence 7, Appli
					Sequence 1, Appli
					Sequence 33, Appl
					Sequence 33, Appl
					Sequence 2803, Ap
					Sequence 3191, Ap
					Sequence 3, Appli
					Sequence 67, Appl
					Sequence 66, Appl
					Sequence 2, Appli
					Sequence 1, Appli

28	39	62.9	419	2	US-09-646-403-4	Sequence 4, Appli
29	39	62.9	419	2	US-09-863-049B-2	Sequence 2, Appli
30	39	62.9	438	2	US-09-198-452A-771	Sequence 771, App
31	39	62.9	835	2	US-09-438-185A-728	Sequence 4768, Ap
32	39	62.9	870	2	US-09-543-681A-4768	Sequence 7268, Ap
33	39	62.9	2733	2	US-09-949-016-11433	Sequence 11433, A
34	39	62.9	3259	2	US-09-949-016-6507	Sequence 6507, Ap
35	38	61.3	20	2	US-09-863-049B-76	Sequence 76, Appl
36	38	61.3	51	2	US-10-094-944-31	Sequence 31, Appl
37	38	61.3	144	2	US-09-270-767-34931	Sequence 34931, A
38	38	61.3	144	2	US-09-270-767-50148	Sequence 50148, A
39	38	61.3	373	2	US-09-248-796A-19926	Sequence 19926, A
40	38	61.3	569	2	US-09-543-681A-6977	Sequence 6977, Ap
41	38	61.3	617	2	US-09-252-991A-23418	Sequence 23418, A
42	38	61.3	932	2	US-08-887-534A-45	Sequence 45, Appl
43	38	61.3	932	2	US-09-527-431-45	Sequence 45, Appl
44	38	61.3	932	2	US-09-446-861-45	Sequence 621, App
45	37	59.7	72	2	US-09-640-211A-621	Sequence 60453, A
46	37	59.7	76	2	US-09-270-767-60453	Sequence 706, App
47	37	59.7	451	2	US-09-538-092-706	Sequence 18, Appl
48	37	59.7	530	2	US-09-861-451A-18	Sequence 20281, A
49	37	59.7	810	2	US-09-248-796A-20281	Sequence 42870, A
50	37	59.7	841	2	US-09-270-767-42870	Sequence 44973, A
51	37	59.7	1024	2	US-09-270-767-44973	Sequence 1, Appli
52	37	59.7	2954	2	US-09-150-867-1	Sequence 4, Appli
53	37	59.7	3457	1	US-08-416-603-4	Sequence 7, Appli
54	36	58.1	148	2	US-09-543-681A-7656	Sequence 7656, Ap
55	36	58.1	153	2	US-09-328-352-6203	Sequence 6203, Ap
56	36	58.1	157	2	US-09-248-796A-26906	Sequence 26906, A
57	36	58.1	245	2	US-09-914-259-65	Sequence 65, Appl
58	36	58.1	245	2	US-10-164-595-34	Sequence 34, Appl
59	36	58.1	281	2	US-09-914-259-63	Sequence 63, Appl
60	36	58.1	294	2	US-09-248-796A-25451	Sequence 25451, A
61	36	58.1	368	2	US-09-248-796A-16492	Sequence 16492, A
62	36	58.1	412	2	US-09-134-000C-6666	Sequence 6666, Ap
63	36	58.1	430	1	US-08-427-993B-7	Sequence 7, Appli
64	36	58.1	430	1	US-08-478-603A-7	Sequence 7, Appli
65	36	58.1	478	2	US-09-134-001C-5065	Sequence 5065, Ap
66	36	58.1	491	2	US-09-029-267-2	Sequence 2, Appli
67	36	58.1	509	1	US-08-427-993B-1	Sequence 1, Appli
68	36	58.1	509	1	US-08-478-609A-1	Sequence 1, Appli
69	36	58.1	560	1	US-08-643-034A-2	Sequence 2, Appli
70	36	58.1	560	2	US-08-648-650A-2	Sequence 2, Appli
71	36	58.1	560	2	US-09-183-266A-2	Sequence 2, Appli
72	36	58.1	588	2	US-09-248-796A-20295	Sequence 20295, A
73	36	58.1	613	2	US-09-446-504-5	Sequence 5, Appli
74	36	58.1	613	2	US-09-712-266-5	Sequence 5, Appli
75	36	58.1	613	2	US-09-091-889A-2	Sequence 26, Appl
76	35	56.5	58	2	US-09-568-051-26	Sequence 23899, A
77	35	56.5	74	2	US-09-248-796A-23899	Sequence 3358, A
78	35	56.5	139	2	US-09-270-767-3358	Sequence 48575, A
79	35	56.5	139	2	US-09-270-767-48575	Sequence 53, Appl
80	35	56.5	139	2	US-09-832-129-53	Sequence 5659, Ap
81	35	56.5	168	2	US-09-107-532A-5659	Sequence 46919, A
82	35	56.5	176	2	US-09-270-767-46919	Sequence 7620, Ap
83	35	56.5	177	2	US-09-543-681A-7620	Sequence 38, Appl
84	35	56.5	230	1	US-09-230-637-38	Sequence 3, Appli
85	35	56.5	230	1	US-08-118-469A-3	Sequence 3, Appli
86	35	56.5	230	2	US-08-909-119-3	Sequence 3, Appli
87	35	56.5	230	2	US-09-152-588-3	Sequence 29, Appl
88	35	56.5	255	2	US-10-094-944-29	Sequence 14945, A
89	35	56.5	285	2	US-09-248-796A-14945	Sequence 7474, Ap
90	35	56.5	298	2	US-09-328-352-7474	Sequence 89, Appl
91	35	56.5	305	2	US-09-818-780-89	Sequence 126, Appl
92	35	56.5	363	2	US-09-716-964B-126	Sequence 34, Appl
93	35	56.5	370	2	US-09-832-129-34	Sequence 4393, Ap
94	35	56.5	370	2	US-09-134-001C-4393	Sequence 3815, Ap
95	35	56.5	371	2	US-09-583-110-3815	Sequence 16527, A
96	35	56.5	378	2	US-09-248-796A-16527	Sequence 4822, Ap
97	35	56.5	378	2	US-09-107-433-4822	Sequence 2250, Ap
98	35	56.5	446	2	US-09-710-279-2250	Sequence 4173, Ap
99	35	56.5	452	2	US-09-134-001C-4173	Sequence 4416, Ap
100	35	56.5	493	2	US-09-107-532A-4416	

101	35	56.5	541	2	US-09-248-796A-17798	Sequence 17798, A	174	34	54.8	683	2	US-09-543-681A-4357	Sequence 4357, Ap
102	35	56.5	601	2	US-09-248-796A-24733	Sequence 24733, A	175	34	54.8	683	2	US-09-538-092-483	Sequence 483, App
103	35	56.5	848	2	US-09-248-796A-18403	Sequence 18403, A	176	34	54.8	792	2	US-09-543-681A-6617	Sequence 6617, Ap
104	35	56.5	905	2	US-09-248-796A-16333	Sequence 16333, A	177	34	54.8	874	2	US-10-163-214-13	Sequence 13, Appl
105	35	56.5	932	2	US-09-248-796A-19128	Sequence 19128, A	178	34	54.8	913	2	US-08-971-089-4	Sequence 4, Appl
106	35	56.5	994	2	US-09-949-016-6779	Sequence 6779, Ap	179	34	54.8	913	2	US-10-117-604A-4	Sequence 4, Appl
107	35	56.5	1055	2	US-09-949-016-6776	Sequence 6776, Ap	180	34	54.8	1025	2	US-09-248-796A-20573	Sequence 20573, A
108	35	56.5	1122	2	US-10-771-708-3	Sequence 3, Appl	181	34	54.8	1401	2	US-08-999-774A-2	Sequence 2, Appl
109	35	56.5	1261	2	US-09-080-855-2	Sequence 2, Appl	182	34	54.8	1401	2	US-09-750-590A-2	Sequence 2, Appl
110	35	56.5	1261	2	US-09-566-076-2	Sequence 2, Appl	183	34	54.8	1967	2	US-09-849-602-16	Sequence 8, Appl
111	35	56.5	1786	2	US-08-973-462-8	Sequence 8, Appl	184	34	54.8	2008	2	US-09-091-501B-8	Sequence 10, Appl
112	35	56.5	1802	2	US-09-949-016-11073	Sequence 11073, A	185	34	54.8	3433	2	US-09-091-501B-10	Sequence 4, Appl
113	35	56.5	2182	1	US-08-487-826B-16	Sequence 16, Appl	186	33.5	54.0	56	2	US-10-127-641-4	Sequence 6, Appl
114	34.5	55.6	582	2	US-09-710-279-1080	Sequence 1080, Ap	187	33.5	54.0	1236	2	US-10-023-649A-4	Sequence 4, Appl
115	34.5	55.6	591	2	US-09-134-001C-3275	Sequence 3275, Ap	188	33.5	54.0	1236	2	US-10-023-649A-6	Sequence 6, Appl
116	34	54.8	22	2	US-09-270-767-60410	Sequence 60410, A	189	33	53.2	20	2	US-09-688-017-195	Sequence 195, App
117	34	54.8	53	1	US-08-676-782-7	Sequence 7, Appl	190	33	53.2	33	2	US-09-491-614B-28	Sequence 28, Appl
118	34	54.8	98	2	US-09-134-000C-6103	Sequence 6103, Ap	191	33	53.2	101	2	US-09-465-901-55	Sequence 55, Appl
119	34	54.8	122	2	US-09-513-999C-8047	Sequence 8047, Ap	192	33	53.2	108	1	US-08-825-782-1	Sequence 1, Appl
120	34	54.8	129	2	US-09-270-767-34216	Sequence 34216, A	193	33	53.2	108	1	US-08-825-782-3	Sequence 3, Appl
121	34	54.8	129	2	US-09-270-767-49433	Sequence 49433, A	194	33	53.2	108	2	US-09-513-999C-7948	Sequence 7948, Ap
122	34	54.8	131	2	US-09-270-767-44935	Sequence 44935, A	195	33	53.2	122	2	US-09-863-766A-4	Sequence 4, Appl
123	34	54.8	134	2	US-09-248-796A-18785	Sequence 18785, A	196	33	53.2	122	2	US-09-583-110-4110	Sequence 4110, Ap
124	34	54.8	139	2	US-09-248-796A-15650	Sequence 15650, A	197	33	53.2	138	2	US-09-902-540-14591	Sequence 14591, A
125	34	54.8	151	2	US-09-248-796A-14713	Sequence 14713, A	198	33	53.2	148	2	US-09-134-000C-5930	Sequence 5930, Ap
126	34	54.8	159	2	US-09-513-999C-5795	Sequence 5795, Ap	199	33	53.2	149	2	US-09-732-210-1369	Sequence 1369, Ap
127	34	54.8	160	2	US-09-134-000C-5539	Sequence 5539, Ap	200	33	53.2	149	2	US-09-107-433-4580	Sequence 4580, Ap
128	34	54.8	179	2	US-09-198-452A-1064	Sequence 1064, Ap	201	33	53.2	150	2	US-09-328-352-7801	Sequence 7801, Ap
129	34	54.8	193	2	US-09-438-185A-920	Sequence 920, App	202	33	53.2	159	2	US-09-328-352-5828	Sequence 5828, Ap
130	34	54.8	193	2	US-09-198-452A-95	Sequence 95, Appl	203	33	53.2	191	2	US-09-248-796A-17226	Sequence 17226, A
131	34	54.8	209	2	US-09-438-185A-70	Sequence 70, Appl	204	33	53.2	197	2	US-09-583-110-4803	Sequence 4803, Ap
132	34	54.8	216	2	US-09-248-796A-23296	Sequence 23296, A	205	33	53.2	202	2	US-09-107-433-3717	Sequence 3717, Ap
133	34	54.8	209	2	US-09-198-452A-627	Sequence 627, App	206	33	53.2	205	2	US-09-710-279-1828	Sequence 1828, Ap
134	34	54.8	223	2	US-09-830-230A-274	Sequence 274, App	207	33	53.2	216	2	US-09-540-236-2088	Sequence 2088, Ap
135	34	54.8	239	2	US-08-851-971-7	Sequence 7, Appl	208	33	53.2	229	2	US-08-936-165A-523	Sequence 523, App
136	34	54.8	241	2	US-09-830-230A-273	Sequence 273, App	209	33	53.2	238	2	US-08-858-207A-420	Sequence 420, App
137	34	54.8	245	2	US-09-914-259-44	Sequence 44, Appl	210	33	53.2	240	2	US-08-851-971-1	Sequence 1, Appl
138	34	54.8	251	2	US-09-902-540-15884	Sequence 15884, A	211	33	53.2	253	2	US-09-543-681A-6581	Sequence 6581, Ap
139	34	54.8	273	2	US-08-235-836C-92	Sequence 92, Appl	212	33	53.2	263	2	US-09-710-279-1418	Sequence 1418, Ap
140	34	54.8	273	2	US-08-235-836C-98	Sequence 98, Appl	213	33	53.2	273	2	US-09-438-185A-204	Sequence 204, App
141	34	54.8	273	2	US-08-235-836C-101	Sequence 101, App	214	33	53.2	276	2	US-09-134-001C-4406	Sequence 4406, Ap
142	34	54.8	273	2	US-08-235-836C-142	Sequence 142, App	215	33	53.2	276	1	US-09-989-386-3	Sequence 3, Appl
143	34	54.8	273	2	US-08-235-836C-144	Sequence 144, App	216	33	53.2	279	2	US-09-543-681A-4548	Sequence 4548, Ap
144	34	54.8	274	2	US-08-235-836C-9	Sequence 9, Appl	217	33	53.2	289	2	US-09-489-039A-12337	Sequence 12337, A
145	34	54.8	277	2	US-08-235-836C-8	Sequence 8, Appl	218	33	53.2	293	2	US-09-138-452A-220	Sequence 220, App
146	34	54.8	281	2	US-09-914-259-45	Sequence 45, Appl	219	33	53.2	294	2	US-09-134-000C-4724	Sequence 4724, Ap
147	34	54.8	296	2	US-09-248-796A-25350	Sequence 25350, A	220	33	53.2	296	2	US-09-944-604-20	Sequence 20, Appl
148	34	54.8	297	2	US-09-710-279-66	Sequence 66, Appl	221	33	53.2	300	2	US-09-328-352-7197	Sequence 1797, Ap
149	34	54.8	297	2	US-09-710-279-746	Sequence 746, App	222	33	53.2	301	2	US-08-944-604-18	Sequence 18, Appl
150	34	54.8	301	2	US-09-248-796A-14357	Sequence 14357, A	223	33	53.2	303	2	US-09-949-016-8260	Sequence 8260, Ap
151	34	54.8	314	2	US-09-640-211A-2179	Sequence 2179, Ap	224	33	53.2	308	2	US-09-949-016-11303	Sequence 11303, A
152	34	54.8	324	2	US-09-134-001C-3365	Sequence 3365, Ap	225	33	53.2	311	2	US-09-248-796A-26549	Sequence 26549, A
153	34	54.8	339	2	US-09-134-000C-4665	Sequence 4665, Ap	226	33	53.2	316	2	US-09-134-000C-4442	Sequence 4442, Ap
154	34	54.8	341	2	US-10-104-047-3424	Sequence 3424, Ap	227	33	53.2	323	1	US-09-019-216-3	Sequence 3, Appl
155	34	54.8	347	2	US-09-902-540-14910	Sequence 14910, A	228	33	53.2	323	2	US-09-249-241-3	Sequence 3, Appl
156	34	54.8	356	2	US-09-252-991A-17157	Sequence 17157, A	229	33	53.2	332	2	US-09-602-787A-122	Sequence 122, App
157	34	54.8	360	2	US-09-640-211A-624	Sequence 624, App	230	33	53.2	338	2	US-09-769-787-161	Sequence 161, App
158	34	54.8	418	2	US-09-359-268A-28	Sequence 28, Appl	231	33	53.2	349	2	US-09-459-774-2	Sequence 2, Appl
159	34	54.8	418	2	US-09-248-796A-18951	Sequence 18951, A	232	33	53.2	349	2	US-09-417-039-7	Sequence 7, Appl
160	34	54.8	451	2	US-09-107-532A-6444	Sequence 6444, Ap	233	33	53.2	349	2	US-09-903-817-2	Sequence 2, Appl
161	34	54.8	481	2	US-09-248-796A-14693	Sequence 14693, A	234	33	53.2	349	2	US-09-107-532A-3670	Sequence 3670, Ap
162	34	54.8	485	2	US-09-328-352-6791	Sequence 6791, Ap	235	33	53.2	350	2	US-09-538-092-111	Sequence 111, App
163	34	54.8	496	2	US-10-600-689-4	Sequence 4, Appl	236	33	53.2	359	2	US-09-248-796A-15803	Sequence 15803, A
164	34	54.8	507	2	US-09-134-000C-4961	Sequence 4961, Ap	237	33	53.2	365	2	US-09-949-016-8710	Sequence 8710, Ap
165	34	54.8	522	2	US-10-104-047-2664	Sequence 2664, Ap	238	33	53.2	372	2	US-09-583-110-2736	Sequence 2736, Ap
166	34	54.8	543	2	US-09-949-016-11672	Sequence 11672, A	239	33	53.2	442	2	US-09-107-433-4227	Sequence 4227, Ap
167	34	54.8	548	2	US-10-104-047-3500	Sequence 3500, Ap	240	33	53.2	445	2	US-09-248-796A-14922	Sequence 14922, A
168	34	54.8	582	2	US-09-919-497-100	Sequence 100, App	241	33	53.2	453	2	US-09-041-075A-5	Sequence 5, Appl
169	34	54.8	620	2	US-09-643-657-13	Sequence 13, Appl	242	33	53.2	480	3	US-08-068-395A-1	Sequence 1, Appl
170	34	54.8	630	2	US-09-248-796A-19619	Sequence 19619, A	243	33	53.2	485	1	US-08-464-365A-1	Sequence 1, Appl
171	34	54.8	660	2	US-09-252-991A-19282	Sequence 19282, A	244	33	53.2	485	1	US-08-068-395A-3	Sequence 3, Appl
172	34	54.8	672	2	US-09-438-185A-587	Sequence 587, App	245	33	53.2	505	1	US-08-464-365-3	Sequence 3, Appl
173	34	54.8	680	2	US-09-252-991A-17566	Sequence 17566, A	246	33	53.2	505	1		

247	33	53.2	510	2	US-09-605-703B-1144	Sequence 1144, Ap	320	52.4	533	4	PCT-US92-00282-3	Sequence 3, Appl
248	33	53.2	522	2	US-09-538-092-1096	Sequence 1096, Ap	321	52.4	534	4	PCT-US92-00282-4	Sequence 4, Appl
249	33	53.2	529	2	US-09-589-733C-16	Sequence 16, Appl	322	52.4	534	2	US-09-949-016-7684	Sequence 7684, Ap
250	33	53.2	529	2	US-09-258-031C-20	Sequence 20, Appl	323	52.4	670	2	US-09-134-001C-2940	Sequence 2940, Ap
251	33	53.2	534	2	US-09-605-703B-1142	Sequence 1142, Ap	324	52.4	712	2	US-09-059-584-51	Sequence 51, Appl
252	33	53.2	540	2	US-09-647-390-18	Sequence 18, Appl	325	52.4	1667	2	US-09-489-039A-11624	Sequence 11624, A
253	33	53.2	540	2	US-09-258-031C-58	Sequence 58, Appl	326	52.4	27	2	US-09-544-684B-22	Sequence 22, Appl
254	33	53.2	576	2	US-09-646-403-8	Sequence 8, Appl	327	51.6	30	2	US-09-486-394-6	Sequence 6, Appl
255	33	53.2	582	2	US-09-702-705-334	Sequence 334, Ap	328	51.6	32	2	US-09-491-614B-9	Sequence 30, Appl
256	33	53.2	582	2	US-09-736-457-334	Sequence 334, Ap	329	51.6	48	2	US-09-515-806-30	Sequence 30, Appl
257	33	53.2	582	2	US-09-614-124B-334	Sequence 334, Ap	330	51.6	56	2	US-09-732-210-1542	Sequence 1542, Ap
258	33	53.2	582	2	US-09-671-325-334	Sequence 334, Ap	331	51.6	62	2	US-09-732-210-1540	Sequence 1540, Ap
259	33	53.2	582	2	US-09-589-184-334	Sequence 334, Ap	332	51.6	82	2	US-08-905-223-481	Sequence 481, Ap
260	33	53.2	582	2	US-09-658-824-334	Sequence 334, Ap	333	51.6	83	2	US-09-513-806-32	Sequence 32, Appl
261	33	53.2	582	2	US-10-017-754-334	Sequence 334, Ap	334	51.6	92	2	US-09-513-999C-6339	Sequence 6339, Ap
262	33	53.2	582	2	US-09-651-563-334	Sequence 334, Ap	335	51.6	94	2	US-09-134-001C-3684	Sequence 3684, Ap
263	33	53.2	582	2	US-09-519-642-334	Sequence 334, Ap	336	51.6	109	2	US-09-328-352-7204	Sequence 7204, Ap
264	33	53.2	583	1	US-08-432-016-2	Sequence 2, Appl	337	51.6	128	2	US-09-107-433-4774	Sequence 4774, Ap
265	33	53.2	583	1	US-08-684-594-2	Sequence 2, Appl	338	51.6	141	2	US-09-270-767-62159	Sequence 62159, A
266	33	53.2	599	2	US-08-556-419-22	Sequence 22, Appl	339	51.6	142	2	US-09-538-092-663	Sequence 663, App
267	33	53.2	604	2	US-09-949-016-9548	Sequence 9548, Ap	340	51.6	143	1	US-07-710-361-9	Sequence 9, Appl
268	33	53.2	605	2	US-09-107-532A-6077	Sequence 6077, Ap	341	51.6	145	2	US-09-621-976-6049	Sequence 6049, Ap
269	33	53.2	629	2	US-08-556-419-23	Sequence 23, Appl	342	51.6	152	1	US-08-318-193-84	Sequence 84, Appl
270	33	53.2	640	2	US-09-248-796A-16129	Sequence 16129, A	343	51.6	156	2	US-09-270-767-46565	Sequence 46565, A
271	33	53.2	640	2	US-09-248-796A-18883	Sequence 18883, A	344	51.6	162	2	US-09-640-211A-610	Sequence 610, App
272	33	53.2	641	2	US-09-270-767-44970	Sequence 44970, A	345	51.6	164	2	US-09-134-001C-5325	Sequence 5325, Ap
273	33	53.2	659	2	US-09-583-110-4043	Sequence 4043, Ap	346	51.6	165	2	US-09-270-767-48569	Sequence 48569, A
274	33	53.2	661	2	US-09-107-433-3700	Sequence 3700, Ap	347	51.6	165	2	US-09-538-092-171	Sequence 171, App
275	33	53.2	679	2	US-09-949-016-7268	Sequence 7268, Ap	348	51.6	177	1	US-08-284-393B-7	Sequence 7, Appl
276	33	53.2	682	2	US-08-836-687B-37	Sequence 37, Appl	349	51.6	177	1	US-08-446-908-4	Sequence 4, Appl
277	33	53.2	697	2	US-09-248-796A-15971	Sequence 15971, A	350	51.6	177	1	US-08-231-205A-4	Sequence 4, Appl
278	33	53.2	713	2	US-09-346-237-9	Sequence 9, Appl	351	51.6	177	1	US-08-871-161-4	Sequence 14, Appl
279	33	53.2	722	2	US-09-248-796A-20613	Sequence 20613, A	352	51.6	177	4	PCT-US95-08950-7	Sequence 7, Appl
280	33	53.2	743	2	US-09-585-858-13	Sequence 13, Appl	353	51.6	177	4	US-08-773-910-3	Sequence 3, Appl
281	33	53.2	743	2	US-10-270-878-13	Sequence 13, Appl	354	51.6	194	1	US-08-408-095-25	Sequence 25, Appl
282	33	53.2	744	2	US-10-270-878-12	Sequence 12, Appl	355	51.6	194	1	US-08-408-095-25	Sequence 3, Appl
283	33	53.2	744	2	US-09-585-858-12	Sequence 12, Appl	356	51.6	194	2	US-09-199-892-3	Sequence 8, Appl
284	33	53.2	828	1	US-08-993-228-21	Sequence 21, Appl	357	51.6	196	2	US-09-513-775B-8	Sequence 10, Appl
285	33	53.2	862	2	US-09-248-796A-16013	Sequence 16013, A	358	51.6	200	2	US-09-949-016-8732	Sequence 8732, Ap
286	33	53.2	868	2	US-09-949-016-9731	Sequence 9731, Ap	359	51.6	201	2	US-09-328-352-4960	Sequence 4960, Ap
287	33	53.2	902	2	US-09-134-000C-6743	Sequence 6743, Ap	360	51.6	202	2	US-09-248-796A-27821	Sequence 27821, A
288	33	53.2	1192	2	US-09-302-540-14526	Sequence 14526, A	361	51.6	209	2	US-09-538-092-500	Sequence 500, App
289	33	53.2	1481	2	US-09-949-016-8693	Sequence 8693, Ap	362	51.6	212	2	US-09-538-092-500	Sequence 35133, A
290	33	53.2	1481	2	US-09-949-016-8694	Sequence 8694, Ap	363	51.6	213	2	US-09-270-767-35133	Sequence 50350, A
291	33	53.2	1481	2	US-09-949-016-8695	Sequence 8695, Ap	364	51.6	213	2	US-09-270-767-50350	Sequence 1, Appl
292	33	53.2	1481	2	US-09-949-016-8696	Sequence 8696, Ap	365	51.6	217	2	US-08-971-158-1	Sequence 3, Appl
293	33	53.2	1486	2	US-09-949-016-8696	Sequence 8696, Ap	366	51.6	217	2	US-08-971-158-3	Sequence 1, Appl
294	33	53.2	1486	2	US-09-949-016-8697	Sequence 8697, Ap	367	51.6	217	2	US-09-416-488-1	Sequence 3, Appl
295	33	53.2	1486	2	US-09-949-016-8677	Sequence 8677, Ap	368	51.6	217	2	US-09-416-488-3	Sequence 2474, Ap
296	33	53.2	1486	2	US-09-949-016-8678	Sequence 8678, Ap	369	51.6	222	2	US-09-710-279-2474	Sequence 2732, Ap
297	33	53.2	1486	2	US-09-949-016-8679	Sequence 8679, Ap	370	51.6	222	2	US-09-710-279-2474	Sequence 4467, Ap
298	33	53.2	1486	2	US-09-949-016-8680	Sequence 8680, Ap	371	51.6	226	2	US-09-134-001C-4467	Sequence 7686, Ap
299	33	53.2	1486	2	US-09-949-016-8681	Sequence 8681, Ap	372	51.6	236	2	US-09-949-016-7686	Sequence 6143, Ap
300	33	53.2	1486	2	US-09-949-016-8682	Sequence 8682, Ap	373	51.6	240	2	US-09-107-532A-6143	Sequence 6143, Ap
301	33	53.2	1499	2	US-09-949-016-8683	Sequence 8683, Ap	374	51.6	241	2	US-09-134-000C-3623	Sequence 3623, Ap
302	33	53.2	1499	2	US-09-949-016-8684	Sequence 8684, Ap	375	51.6	252	2	US-09-538-092-198	Sequence 198, App
303	33	53.2	1499	2	US-09-949-016-8685	Sequence 8685, Ap	376	51.6	265	2	US-09-252-991A-28073	Sequence 28073, A
304	33	53.2	1499	2	US-09-949-016-8686	Sequence 8686, Ap	377	51.6	271	2	US-09-248-796A-27391	Sequence 27391, A
305	33	53.2	1499	2	US-09-949-016-8687	Sequence 8687, Ap	378	51.6	273	2	US-09-270-767-4062	Sequence 44062, A
306	33	53.2	1499	2	US-09-949-016-8688	Sequence 8688, Ap	379	51.6	280	2	US-09-543-684A-4828	Sequence 4828, Ap
307	33	53.2	1499	2	US-09-949-016-8689	Sequence 8689, Ap	380	51.6	284	2	US-09-134-000C-3648	Sequence 3648, Ap
308	33	53.2	1499	2	US-09-949-016-8690	Sequence 8690, Ap	381	51.6	285	1	US-08-149-809-24	Sequence 24, Appl
309	33	53.2	1499	2	US-09-949-016-8691	Sequence 8691, Ap	382	51.6	295	2	US-09-248-796A-18333	Sequence 18333, A
310	33	53.2	2349	2	US-09-538-092-914	Sequence 914, App	383	51.6	311	2	US-09-134-001C-4235	Sequence 4235, Ap
311	33	53.2	5024	2	US-09-710-279-2964	Sequence 2964, Ap	384	51.6	316	2	US-09-540-236-2499	Sequence 2499, Ap
312	33	53.2	10182	2	US-09-134-001C-3159	Sequence 3159, Ap	385	51.6	325	2	US-09-328-352-4730	Sequence 4730, Ap
313	32.5	52.4	74	4	PCT-US92-00282-24	Sequence 24, Appl	386	51.6	341	2	US-09-489-039A-9238	Sequence 9238, Ap
314	32.5	52.4	109	2	US-09-248-796A-18216	Sequence 18216, A	387	51.6	341	2	US-09-489-039A-10987	Sequence 10987, A
315	32.5	52.4	245	2	US-09-305-856B-18	Sequence 18, Appl	388	51.6	341	2	US-09-489-039A-11246	Sequence 11246, A
316	32.5	52.4	299	2	US-09-270-767-46195	Sequence 46195, A	389	51.6	341	2	US-09-489-039A-11507	Sequence 11507, A
317	32.5	52.4	400	2	US-09-248-796A-18122	Sequence 18122, A	390	51.6	343	2	US-09-328-352-4130	Sequence 4130, Ap
318	32.5	52.4	531	4	PCT-US92-00282-5	Sequence 5, Appl	391	51.6	343	2	US-09-489-039A-11316	Sequence 11316, A
319	32.5	52.4	533	2	US-09-949-016-5947	Sequence 5947, Ap	392	51.6	347	2	US-09-538-092-753	Sequence 753, App

```
393 32 51.6 360 2 US-09-248-796A-20433 Sequence 20433, A Sequence 19, Appl
394 32 51.6 364 2 US-09-489-039A-8104 Sequence 8104, Ap Sequence 19, Appl
395 32 51.6 367 2 US-10-094-944-21 Sequence 21, Ap Sequence 5, Appl
396 32 51.6 387 2 US-09-248-796A-18228 Sequence 18228, A Sequence 19, Appl
397 32 51.6 395 2 US-09-328-352-7437 Sequence 7437, Ap Sequence 19, Appl
398 32 51.6 399 2 US-09-323-872A-17 Sequence 17, Appl Sequence 19, Appl
399 32 51.6 399 2 US-09-072-433-25 Sequence 25, Appl Sequence 6, Appl
400 32 51.6 402 2 US-09-107-532A-5693 Sequence 5693, Ap Sequence 19, Appl
401 32 51.6 406 2 US-09-248-796A-20860 Sequence 20860, A Sequence 19, Appl
402 32 51.6 424 2 US-09-491-577-88 Sequence 88, Ap Sequence 19, Appl
403 32 51.6 434 2 US-09-248-796A-18389 Sequence 18389, A Sequence 736, App
404 32 51.6 438 2 US-09-248-796A-19338 Sequence 19338, A Sequence 1668, Ap
405 32 51.6 452 2 US-09-711-164-366 Sequence 366, App Sequence 11, Appl
406 32 51.6 474 2 US-09-248-796A-16728 Sequence 16728, A Sequence 16, Appl
407 32 51.6 498 2 US-09-238-303-8 Sequence 8, Appl Sequence 6999, Ap
408 32 51.6 498 2 US-09-946-239-8 Sequence 8, Appl Sequence 7880, Ap
409 32 51.6 501 2 US-10-104-047-3448 Sequence 3448, Ap Sequence 2, Appl
410 32 51.6 505 2 US-09-134-000C-5340 Sequence 5340, Ap Sequence 8, Appl
411 32 51.6 513 2 US-09-489-039A-14224 Sequence 14224, A Sequence 23, Appl
412 32 51.6 515 2 US-10-104-047-2411 Sequence 2411, Ap Sequence 19, Appl
413 32 51.6 516 2 US-09-252-991A-29719 Sequence 29719, A Sequence 10918, A
414 32 51.6 522 2 US-09-949-016-7549 Sequence 7549, Ap Sequence 7, Appl
415 32 51.6 529 2 US-09-589-733C-17 Sequence 17, Appl Sequence 7, Appl
416 32 51.6 529 2 US-09-647-390-16 Sequence 16, Appl Sequence 8, Appl
417 32 51.6 538 2 US-09-258-031C-16 Sequence 16, Appl Sequence 8, Appl
418 32 51.6 545 2 US-09-270-767-39221 Sequence 39221, A Sequence 5821, Ap
419 32 51.6 545 2 US-09-270-767-54438 Sequence 54438, A Sequence 1, Appl
420 32 51.6 563 2 US-09-540-236-2939 Sequence 2939, Ap Sequence 1136, Ap
421 32 51.6 572 2 US-09-248-796A-19066 Sequence 19066, A Sequence 9908, Ap
422 32 51.6 579 2 US-09-538-092-1284 Sequence 1284, Ap Sequence 5, Appl
423 32 51.6 592 2 US-09-270-767-42535 Sequence 42535, A Sequence 1262, Ap
424 32 51.6 603 2 US-09-248-796A-17178 Sequence 17178, A Sequence 2269, Ap
425 32 51.6 609 2 US-09-328-352-8168 Sequence 8168, Ap Sequence 10, Appl
426 32 51.6 609 2 US-09-107-532A-4451 Sequence 4451, Ap Sequence 10, Appl
427 32 51.6 619 2 US-09-578-441-3 Sequence 3, Appl Sequence 10, Appl
428 32 51.6 620 2 US-09-578-441-4 Sequence 4, Appl Sequence 10, Appl
429 32 51.6 654 2 US-09-902-540-12072 Sequence 12072, A Sequence 10, Appl
430 32 51.6 714 2 US-09-949-016-10296 Sequence 10296, A Sequence 10, Appl
431 32 51.6 737 2 US-09-949-016-9368 Sequence 9368, Ap Sequence 10, Appl
432 32 51.6 774 2 US-09-328-352-7534 Sequence 7534, Ap Sequence 10, Appl
433 32 51.6 810 2 US-09-902-540-11337 Sequence 11337, A Sequence 10, Appl
434 32 51.6 816 2 US-09-489-039A-10520 Sequence 10520, A Sequence 10, Appl
435 32 51.6 868 2 US-09-538-092-787 Sequence 787, App Sequence 10, Appl
436 32 51.6 943 2 US-09-949-016-6758 Sequence 6758, Ap Sequence 11702, A
437 32 51.6 955 2 US-09-949-016-10122 Sequence 10122, A Sequence 9991, Ap
438 32 51.6 967 2 US-09-914-259-21 Sequence 21, Appl Sequence 9, Appl
439 32 51.6 1092 2 US-09-538-092-122 Sequence 122, App Sequence 9, Appl
440 32 51.6 1128 1 US-08-111-939-2 Sequence 2, Appl Sequence 9, Appl
441 32 51.6 1128 2 US-09-641-741-30 Sequence 30, Appl Sequence 9, Appl
442 32 51.6 1128 2 US-09-060-482-8 Sequence 8, Appl Sequence 1044, Ap
443 32 51.6 1208 2 US-09-489-039A-8724 Sequence 8724, Ap Sequence 10, Appl
444 32 51.6 1226 2 US-09-601-537-7 Sequence 7, Appl Sequence 10, Appl
445 32 51.6 1227 1 US-08-760-075A-18 Sequence 18, Appl Sequence 10, Appl
446 32 51.6 1227 1 US-09-338-546-18 Sequence 18, Appl Sequence 10, Appl
447 32 51.6 1227 2 US-09-659-084-18 Sequence 18, Appl Sequence 10, Appl
448 32 51.6 1244 2 US-09-949-016-11702 Sequence 11702, A Sequence 10, Appl
449 32 51.6 1256 2 US-09-949-016-9991 Sequence 9991, Ap Sequence 10, Appl
450 32 51.6 1296 2 US-08-857-636-60 Sequence 60, Appl Sequence 10, Appl
451 32 51.6 1313 2 US-08-989-299-9 Sequence 9, Appl Sequence 10, Appl
452 32 51.6 1313 2 US-09-407-427-9 Sequence 9, Appl Sequence 10, Appl
453 32 51.6 1313 2 US-09-635-501-9 Sequence 9, Appl Sequence 1044, Ap
454 32 51.6 1424 2 US-09-538-092-1044 Sequence 1044, Ap Sequence 10, Appl
455 32 51.6 1434 1 US-08-540-406-10 Sequence 10, Appl Sequence 10, Appl
456 32 51.6 1434 2 US-08-656-055-10 Sequence 10, Appl Sequence 10, Appl
457 32 51.6 1434 2 US-08-954-668-10 Sequence 10, Appl Sequence 10, Appl
458 32 51.6 1434 2 US-08-918-658-10 Sequence 10, Appl Sequence 10, Appl
459 32 51.6 1434 2 US-09-724-631-10 Sequence 10, Appl Sequence 10, Appl
460 32 51.6 1434 2 US-08-954-701A-10 Sequence 10, Appl Sequence 10, Appl
461 32 51.6 1434 2 US-09-754-032-10 Sequence 10, Appl Sequence 10, Appl
462 32 51.6 1434 2 US-08-916-140-10 Sequence 10, Appl Sequence 10, Appl
463 32 51.6 1434 4 PCT-US95-13233-10 Sequence 10, Appl Sequence 10, Appl
464 32 51.6 1438 2 US-09-538-092-189 Sequence 189, App Sequence 19, Appl
465 32 51.6 1447 1 US-08-540-406-19 Sequence 19, Appl Sequence 19, Appl
```

ALIGNMENTS

RESULT 1

```
US-09-949-016-8330
; Sequence 8330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8330
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-8330
```

Query Match 100.0%; Score 62; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

DB 478 ILEENKELENALK 490

RESULT 2
US-08-230-491A-2
; Sequence 2, Application US/08230491A
; Patent No. 5587299
; GENERAL INFORMATION:
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
; FIBROBLAST ACTIVATION PROTEIN AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FELPE & LYNCH
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT - ASC II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,491A
; FILING DATE: 20-APRIL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587299man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-230-491A-2

Query Match 100.0%; Score 62; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||||
DB 487 ILEENKELENALK 499

RESULT 3
US-08-619-280A-2
; Sequence 2, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; FIBROBLAST ACTIVATION PROTEIN AND USES THEREOF
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felpe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-619-280A-2

Query Match 100.0%; Score 62; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||||
DB 487 ILEENKELENALK 499

RESULT 4
US-08-940-391-2
; Sequence 2, Application US/08940391
; Patent No. 5965373
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
; FIBROBLAST ACTIVATION PROTEIN AND USES THEREOF
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felpe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

US-08-940-391-2

Query Match 100.0%; Score 62; DB 1; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
|||:|||||
Db 487 ILEENKELENALK 499

RESULT 5

US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:

CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-2

Query Match 100.0%; Score 62; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
|||:|||||
Db 487 ILEENKELENALK 499

RESULT 6

US-09-518-550-26
; Sequence 26, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
; APPLICANT: TRAVIS, James
; APPLICANT: POTEMPA, Jan
; APPLICANT: BANBULA, Agnieszka

; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
; FILE REFERENCE: 235.00190101
; CURRENT APPLICATION NUMBER: US/09/518,550
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,148
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US00/05551
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 761
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-518-550-26

Query Match 88.7%; Score 55; DB 2; Length 761;
Best Local Similarity 76.9%; Pred. No. 1.3;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
:|||||:|
Db 487 VLEENKELENSLR 499

RESULT 7

US-10-012-819-82
; Sequence 82, Application US/10012819
; Patent No. 6916615
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-012-819-82

Query Match 69.4%; Score 43; DB 2; Length 346;
Best Local Similarity 69.2%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
|||:|||||
Db 157 ILTESKIEEALK 169

RESULT 8

US-09-538-092-649
; Sequence 649, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387

```

; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 649
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YNL021W
US-09-538-092-649

Query Match          69.4%; Score 43; DB 2; Length 706;
Best Local Similarity 81.8%; Pred. No. 90;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LEENKELENAL 12
Db      22 LEENKEENSLS 32
      : ||||| ||| :
      : ||||| ||| :

RESULT 9
US-09-949-002-482
; Sequence 482, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Human
US-09-949-002-482

Query Match          67.7%; Score 42; DB 2; Length 825;
Best Local Similarity 61.5%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
Db      401 VTEENKELANELR 413
      : ||||| ||| :
      : ||||| ||| :

RESULT 10
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match          67.7%; Score 42; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
Db      301 VTEENKELANELR 313
      : ||||| ||| :
      : ||||| ||| :

RESULT 11
US-09-657-472-2
; Sequence 2, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolik, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/09/657,472
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-657-472-2

Query Match          67.7%; Score 42; DB 2; Length 1170;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
Db      301 VTEENKELANELR 313
      : ||||| ||| :
      : ||||| ||| :

RESULT 12
US-09-949-002-350
; Sequence 350, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```


;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;; FILE OF INVENTION: AND USES THEREOF
;; FILE REFERENCE: CL000790
;; CURRENT APPLICATION NUMBER: US/09/949,002
;; CURRENT FILING DATE: 2000-01-28
;; PRIOR APPLICATION NUMBER: 60/231,401
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 10823
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 350
;; LENGTH: 1170
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-002-350

Query Match 67.7%; Score 42; DB 2; Length 1170;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313

RESULT 13
US-09-270-767-42366
Sequence 42366, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42366
LENGTH: 317
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-42366

Query Match 66.1%; Score 41; DB 2; Length 317;
Best Local Similarity 61.5%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 62 LLERNKELETFFIK 74

RESULT 14
US-09-180-167A-7
Sequence 7, Application US/09180167A
Patent No. 6558950
GENERAL INFORMATION:
APPLICANT: Gordon C. Shore et al.
TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
FILE REFERENCE: 50013/004003
CURRENT APPLICATION NUMBER: US/09/180,167A
CURRENT FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: CA 2,198,988
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: PCT/IB98/00706
PRIOR FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 62
TYPE: PRT
ORGANISM: Mus musculus

Query Match 64.5%; Score 40; DB 2; Length 62;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

US-09-180-167A-7

Query Match 64.5%; Score 40; DB 2; Length 62;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
Db 9 LEENKSLKNDLR 20

RESULT 15
US-09-033-524B-7
Sequence 7, Application US/09033524B
Patent No. 6607880
GENERAL INFORMATION:
APPLICANT: Gordon C. Shore et al.
TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
FILE REFERENCE: 50013/004002
CURRENT APPLICATION NUMBER: US/09/033,524B
CURRENT FILING DATE: 1998-03-02
PRIOR APPLICATION NUMBER: CA 2,198,988
PRIOR FILING DATE: 1997-03-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 62
TYPE: PRT
ORGANISM: Mus musculus
US-09-033-524B-7

Query Match 64.5%; Score 40; DB 2; Length 62;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
Db 9 LEENKSLKNDLR 20

RESULT 16
US-08-985-526-1
Sequence 1, Application US/08985526
Patent No. 6080728
GENERAL INFORMATION:
APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
FILE REFERENCE: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorow Jr., Robert G
TELECOMMUNICATION INFORMATION:

Query Match 64.5%; Score 40; DB 2; Length 62;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
Db 9 LEENKSLKNDLR 20

TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-985-526-1

Query Match 64.5%; Score 40; DB 2; Length 218;
Best Local Similarity 72.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 BENKELENALK 13
||||| |:
DB 3 BENKELANELR 13

RESULT 17
US-09-180-167A-33
Sequence 33, Application US/09180167A
Patent No. 6558950
GENERAL INFORMATION:
APPLICANT: Gordon C. Shore et al.
TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
FILE REFERENCE: 50013/004003
CURRENT APPLICATION NUMBER: US/09/180,167A
CURRENT FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: CA 2,198,988
PRIOR FILING DATE: 1997-03-03
PRIOR APPLICATION NUMBER: PCT/IB98/00706
PRIOR FILING DATE: 1998-03-02
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 245
TYPE: PRT
ORGANISM: Mus musculus
US-09-180-167A-33

Query Match 64.5%; Score 40; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
||||| |:
DB 177 LEENKSLKNDLR 188

RESULT 18
US-09-033-524B-33
Sequence 33, Application US/09033524B
Patent No. 6607880
GENERAL INFORMATION:
APPLICANT: Gordon C. Shore et al.
TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING
FILE REFERENCE: 50013/004002
CURRENT APPLICATION NUMBER: US/09/033,524B
CURRENT FILING DATE: 1998-03-02
PRIOR APPLICATION NUMBER: CA 2,198,988
PRIOR FILING DATE: 1997-03-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 245
TYPE: PRT
ORGANISM: Mus musculus
US-09-033-524B-33

Query Match 64.5%; Score 40; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 84;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 LEENKELENALK 13
||||| |:
DB 177 LEENKSLKNDLR 188

RESULT 19
US-08-097-831-2
Sequence 2, Application US/08097831
Patent No. 5510473
GENERAL INFORMATION:
APPLICANT: Camerini-Otero, Rafael D.
APPLICANT: Angov, Evangelina
TITLE OF INVENTION: Cloning and Expression of Tag reca
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,831
FILING DATE: 19930726
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: NIH066.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-097-831-2

Query Match 64.5%; Score 40; DB 1; Length 340;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 2 LEEN--KELENALK 13
:|||| |:
DB 1 MEENKRSLENALK 14

RESULT 20
US-09-583-110-2803
Sequence 2803, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02

```
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2803
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2803

Query Match 64.5%; Score 40; DB 2; Length 391;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 2 LEENKE--LENALK 13
||||:|:||||
Db 269 LEENREKSLKALK 282

RESULT 21
US-09-107-433-3191
; Sequence 3191, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3191:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...403
; SEQUENCE DESCRIPTION: SEQ ID NO: 3191:
US-09-107-433-3191

Query Match 64.5%; Score 40; DB 2; Length 403;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 2 LEENKE--LENALK 13
```

```
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2803
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2803

Query Match 64.5%; Score 40; DB 2; Length 391;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 2 LEENKE--LENALK 13
||||:|:||||
Db 269 LEENREKSLKALK 282

RESULT 22
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-985-526-3

Query Match 64.5%; Score 40; DB 2; Length 441;
Best Local Similarity 72.7%; Pred. No. 1.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EENKELENALK 13
||||:|:||||
Db 3 EENKELENALK 13

RESULT 23
US-09-914-259-67
; Sequence 67, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 959
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-914-259-67
```

```

Query Match      64.5%; Score 40; DB 2; Length 959;
Best Local Similarity 53.8%; Pred. No. 3.7e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
       :|:|||||:|:|
Db      853 LLQETKELKNEIK 865

RESULT 24
US-09-914-259-66
; Sequence 66, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-914-259-66

Query Match      64.5%; Score 40; DB 2; Length 961;
Best Local Similarity 53.8%; Pred. No. 3.7e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
       :|:|||||:|:|
Db      852 LLQETKELKNEIK 864

RESULT 25
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match      64.5%; Score 40; DB 2; Length 2285;
Best Local Similarity 61.5%; Pred. No. 9.5e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
       :|:|||||:|:|
Db      847 ILKTNKELEKAE 859

RESULT 26
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. 6833265
; GENERAL INFORMATION:

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-646-403-4

Query Match 62.9%; Score 39; DB 2; Length 419;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|||||:|:|:|:
DB 55 LEENQELRDAIR 66

RESULT 29
US-09-863-049B-2
; Sequence 2, Application US/09863049B
; Patent No. 6824972
; GENERAL INFORMATION:
; APPLICANT: Kenrick, Sue J.
; APPLICANT: Nelson, David L.
; APPLICANT: Aradhya, Swaroop
; APPLICANT: D'Urso, Michele
; APPLICANT: Woffendin, Hayley
; APPLICANT: Munnich, Arnold
; APPLICANT: Smahi, Asmaa
; APPLICANT: Israel, Alain
; APPLICANT: Poustka, Annemarie
; APPLICANT: Lewis, Richard A
; APPLICANT: Levy, Moise
; APPLICANT: Heiss, Nina
; TITLE OF INVENTION: Diagnosis and Treatment of Medical Conditions Associated with Def
; FILE REFERENCE: HO-P01961US1
; CURRENT APPLICATION NUMBER: US/09/863,049B
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/206,223
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: human
US-09-863-049B-2

Query Match 62.9%; Score 39; DB 2; Length 419;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|||||:|:|:|:
DB 55 LEENQELRDAIR 66

RESULT 30
US-09-198-452A-771
; Sequence 771, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 771
; LENGTH: 438

; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-771

Query Match 62.9%; Score 39; DB 2; Length 438;
Best Local Similarity 72.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENA 11
|||||:|:|:|:
DB 42 ILEEAKELDNS 52

RESULT 31
US-09-438-185A-728
; Sequence 728, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; TITLE OF INVENTION: The Regents of the University of California
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 728
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPh0726
US-09-438-185A-728

Query Match 62.9%; Score 39; DB 2; Length 835;
Best Local Similarity 72.7%; Pred. No. 4.5e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENA 11
|||||:|:|:|:
DB 439 ILEEAKELDNS 449

RESULT 32
US-09-543-681A-4768
; Sequence 4768, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4768
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4768

Query Match 62.9%; Score 39; DB 2; Length 870;
Best Local Similarity 72.7%; Pred. No. 4.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENAL 12
|||||
Db 754 LEENIELENTI 764

RESULT 33

US-09-949-016-11433
; Sequence 11433, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11433
; LENGTH: 2733
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11433

Query Match 62.9%; Score 39; DB 2; Length 2733;
Best Local Similarity 77.8%; Pred. No. 1.6e+03;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LEENKELEN 10
|:::|||||
Db 2033 LQONKELEN 2041

RESULT 34

US-09-949-016-6507
; Sequence 6507, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6507
; LENGTH: 3259
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6507

Query Match 62.9%; Score 39; DB 2; Length 3259;
Best Local Similarity 77.8%; Pred. No. 2e+03;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LEENKELEN 10
|:::|||||
Db 2559 LQONKELEN 2567

RESULT 35

US-09-863-049B-76
; Sequence 76, Application US/09863049B
; Patent No. 6824972

GENERAL INFORMATION:

; APPLICANT: Kenrick, Sue J.
; APPLICANT: Nelson, David L.
; APPLICANT: Arachya, Swaroop
; APPLICANT: D'Urso, Michele
; APPLICANT: Woffendin, Hayley
; APPLICANT: Munnich, Arnold
; APPLICANT: Smahi, Asmae
; APPLICANT: Israel, Alain
; APPLICANT: Poustka, Annemarie
; APPLICANT: Lewis, Richard A
; APPLICANT: Levy, Moise
; APPLICANT: Heiss, Nina
; TITLE OF INVENTION: Diagnosis and Treatment of Medical Conditions Associated with Def
; FILE REFERENCE: HO-P01961US1
; CURRENT APPLICATION NUMBER: US/09/863,049B
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/206,223
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Human
US-09-863-049B-76

Query Match 61.3%; Score 38; DB 2; Length 20;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|:::|:::
Db 5 LEENQELPAIR 16

RESULT 36

US-10-094-944-31
; Sequence 31, Application US/10094944
; Patent No. 6958387

GENERAL INFORMATION:

; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: DuBose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/10/094,944
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-944-31

Query Match 61.3%; Score 38; DB 2; Length 51;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENAL 12
|:::|:::
Db 31 LEENRSMESAL 41

RESULT 37

US-09-270-767-34931
; Sequence 34931, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34931
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-34931

Query Match 61.3%; Score 38; DB 2; Length 144;
Best Local Similarity 63.6%; Pred. No. 96;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEENKELENAL 12
|||:::|
Db 98 LEENKKLODAV 108

RESULT 38
US-09-270-767-50148
; Sequence 50148, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50148
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-50148

Query Match 61.3%; Score 38; DB 2; Length 144;
Best Local Similarity 63.6%; Pred. No. 96;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEENKELENAL 12
|||:::|
Db 98 LEENKKLODAV 108

RESULT 39
US-09-248-796A-19926
; Sequence 19926, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19926
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Candida albicans

US-09-248-796A-19926

Query Match 61.3%; Score 38; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 KELENALK 13
|||||||
Db 143 KELENALK 150

RESULT 40

US-09-543-681A-6977
; Sequence 6977, Application US/09543681A
; Patent No. 6605709; GENERAL INFORMATION:
; APPLICANT: GARY BRETON; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344; SEQ ID NO 6977
; LENGTH: 569; TYPE: PRT
; ORGANISM: Proteus mirabilis

US-09-543-681A-6977

Query Match 61.3%; Score 38; DB 2; Length 569;
Best Local Similarity 72.7%; Pred. No. 4.3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENAL 12
|||:::|
Db 338 LKENGELNTL 348

RESULT 41

US-09-252-991A-23418

; Sequence 23418, Application US/09252991A
; Patent No. 6551795; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142; SEQ ID NO 23418
; LENGTH: 617; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23418

Query Match 61.3%; Score 38; DB 2; Length 617;
Best Local Similarity 63.6%; Pred. No. 4.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENA 11
|||:::|
Db 340 VLERTRELENA 350

RESULT 42

US-08-887-534A-45

Sequence 45, Application US/08887534A
Patent No. 6455323
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-527-431-45
Query Match 61.3%; Score 38; DB 2; Length 932;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 LEENKELENALK 13
Db 919 LVQNKIENALK 930
RESULT 44
US-09-446-861-45
Sequence 45, Application US/09446861
Patent No. 6740485
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Li-Hsien Rin-Laures, M.D.
STREET: 6300 Sears Tower 233 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/446,861
FILING DATE: 22-Mar-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Li-Hsien Rin-Laures, M.D.
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/6314.PCP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-446-861-45
Query Match 61.3%; Score 38; DB 2; Length 932;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 LEENKELENALK 13
Db 919 LVQNKIENALK 930

Sequence 45, Application US/08887534A
Patent No. 6455323
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-887-534A-45
Query Match 61.3%; Score 38; DB 2; Length 932;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 LEENKELENALK 13
Db 919 LVQNKIENALK 930
RESULT 43
US-09-527-431-45
Sequence 45, Application US/09527431
Patent No. 6485899
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/527,431
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534


```
RESULT 45
US-09-640-211A-621
; Sequence 621, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 621
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-621

Query Match          59.7%; Score 37; DB 2; Length 72;
Best Local Similarity 66.7%; Pred. No. 65;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy  2 LEENKELENALK 13
Db  7 LEKNFELENKLE 18

RESULT 46
US-09-270-767-60453
; Sequence 60453, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60453
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60453

Query Match          59.7%; Score 37; DB 2; Length 76;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  4 ENKELENALK 13
Db  47 ENEELWNALK 56

RESULT 47
US-09-538-092-706
; Sequence 706, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1999-02-12
; CURRENT APPLICATION NUMBER: US/09/248,796A
; FILE REFERENCE: 107196.132
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; APPLICANT: Keith Weinstein et al
; GENERAL INFORMATION:
; Sequence 20281, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

```
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 706
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YOL069W
US-09-538-092-706

Query Match          59.7%; Score 37; DB 2; Length 451;
Best Local Similarity 72.7%; Pred. No. 4.8e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  3 EENKELENALK 13
Db  190 EQNKELEIRQLK 200

RESULT 48
US-09-861-451A-18
; Sequence 18, Application US/09861451A
; Patent No. 6759516
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: PF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PP7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD662
US-09-861-451A-18

Query Match          59.7%; Score 37; DB 2; Length 530;
Best Local Similarity 61.5%; Pred. No. 5.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy  1 ILEENKELENALK 13
Db  384 ILNKNKEIRINAQK 396

RESULT 49
US-09-248-796A-20281
; Sequence 20281, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

SEQ ID NO 20281
LENGTH: 810
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-20281

Query Match 59.7%; Score 37; DB 2; Length 810;
Best Local Similarity 53.8%; Pred. No. 9e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|:|||||:
DB 755 ILENKLERELE 767

RESULT 50
US-09-270-767-42870
Sequence 42870, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42870
LENGTH: 841
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42870

Query Match 59.7%; Score 37; DB 2; Length 841;
Best Local Similarity 66.7%; Pred. No. 9.4e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENAL 12
|:|||||:
DB 744 ILEENENQENPL 755

RESULT 51
US-09-270-767-44973
Sequence 44973, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 44973
LENGTH: 1024
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44973

Query Match 59.7%; Score 37; DB 2; Length 1024;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ENKELENALK 13
|:|||||:
DB 995 ENEELNALK 1004

RESULT 52
US-09-150-867-1
Sequence 1, Application US/09150867
Patent No. 6645748
GENERAL INFORMATION:

APPLICANT: Wood, Kenneth W.
APPLICANT: Sakowicz, Roman
APPLICANT: Goldstein, Lawrence S.B.
APPLICANT: Cleveland, Don W.
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Plus End-Directed Microtubule Motor Required for
TITLE OF INVENTION: Chromosome Congression
FILE REFERENCE: 18557C-000110US
CURRENT APPLICATION NUMBER: US/09/150,867
CURRENT FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: US 60/058,645
EARLIER FILING DATE: 1997-09-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2954
TYPE: PRT
ORGANISM: Xenopus sp.
FEATURE:
OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E)
OTHER INFORMATION: member of the kinesin superfamily of microtubule
OTHER INFORMATION: motor proteins
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(472)
OTHER INFORMATION: kinesin like motor domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (473)..(2752)
OTHER INFORMATION: rod domain
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (2753)..(2954)
OTHER INFORMATION: tail domain
US-09-150-867-1

Query Match 59.7%; Score 37; DB 2; Length 2954;
Best Local Similarity 63.6%; Pred. No. 3.7e+03;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENA 11
|:|||||:
DB 1734 LLENKELEQS 1744

RESULT 53
US-08-416-603-4
Sequence 4, Application US/08416603
Patent No. 5866780
GENERAL INFORMATION:

APPLICANT: Law, Marcus
APPLICANT: Hebara, Ledare
APPLICANT: Reddick, Bradford B.
TITLE OF INVENTION: Maize Chlorotic Dwarf Virus Genome and
TITLE OF INVENTION: Uses Therefor
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,603
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lloyd, Jeffrey
 ; REGISTRATION NUMBER: 35,589
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3457 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-416-603-4

Query Match 59.7%; Score 37; DB 1; Length 3457;
 Best Local Similarity 66.7%; Pred. No. 4.3e+03;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
 |||||: ||:
 Db 2257 LERNKELKALE 2268

RESULT 54
 US-09-543-681A-7656
 ; Sequence 7656, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 7656
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 ; US-09-543-681A-7656

Query Match 58.1%; Score 36; DB 2; Length 148;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENAL 12
 : |||||
 Db 1 VLMEQLELENAL 12

RESULT 55
 US-09-328-352-6203
 ; Sequence 6203, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: GTC99-03PA
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 6203
 ; LENGTH: 153
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-6203

Query Match 58.1%; Score 36; DB 2; Length 153;
 Best Local Similarity 61.5%; Pred. No. 2.1e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
 : |||||: ||:
 Db 33 LTERKELFDALK 45

RESULT 56
 US-09-248-796A-26906
 ; Sequence 26906, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 26906
 ; LENGTH: 157
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-26906

Query Match 58.1%; Score 36; DB 2; Length 157;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NKELENALK 13
 ||::|||:
 Db 18 NKKIENAIK 26

RESULT 57
 US-09-914-259-65
 ; Sequence 65, Application US/09914259
 ; Patent No. 6495336
 ; GENERAL INFORMATION:
 ; APPLICANT: Makowski, Lee
 ; APPLICANT: Hyman, Paul
 ; APPLICANT: Williams, Mark
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 ; FILE REFERENCE: 8471-010-999
 ; CURRENT APPLICATION NUMBER: US/09/914,259
 ; CURRENT FILING DATE: 2000-11-21
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 65
 ; LENGTH: 245
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-914-259-65

Query Match 58.1%; Score 36; DB 2; Length 245;
 Best Local Similarity 59.3%; Pred. No. 3.5e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
 ||::|||:
 Db 228 LEQNRRLTNELK 239

RESULT 58
 US-10-164-595-34
 ; Sequence 34, Application US/10164595
 ; Patent No. 6657054

GENERAL INFORMATION:
 APPLICANT: Origene Technologies, Inc
 TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
 FILE REFERENCE: IU 103 R1
 CURRENT APPLICATION NUMBER: US/10/164,595
 CURRENT FILING DATE: 2002-06-10
 NUMBER OF SEQ ID NOS: 80
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 34
 LENGTH: 245
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-164-595-34

Query Match 58.1%; Score 36; DB 2; Length 245;
 Best Local Similarity 58.3%; Pred. No. 3.5e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LEENKELENALK 13
 Db 228 LEQNRLTNELK 239

RESULT 59
 US-09-914-259-63
 Sequence 63, Application US/09914259 -
 Patent No. 6495336
 GENERAL INFORMATION:
 APPLICANT: Makowski, Lee
 APPLICANT: Hyman, Paul
 APPLICANT: Williams, Mark
 TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 FILE REFERENCE: 8471-010-999
 CURRENT APPLICATION NUMBER: US/09/914,259
 CURRENT FILING DATE: 2000-11-21
 NUMBER OF SEQ ID NOS: 180
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 63
 LENGTH: 281
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-09-914-259-63

Query Match 58.1%; Score 36; DB 2; Length 281;
 Best Local Similarity 58.3%; Pred. No. 4.1e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LEENKELENALK 13
 Db 264 LEQNRLTNELK 275

RESULT 60
 US-09-248-796A-25451
 Sequence 25451, Application US/09248796A
 Patent No. 6747137
 GENERAL INFORMATION:
 APPLICANT: Keith Weinstock et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248,796A
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 25451
 LENGTH: 294
 TYPE: PRT
 ORGANISM: Candida albicans
 US-09-248-796A-25451

Query Match 58.1%; Score 36; DB 2; Length 294;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LEENKELENALK 13
 Db 84 IEDNKENQNGVK 95
 Search completed: March 11, 2006, 12:27:22
 Job time : 34.0789 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:26:08 ; Search time 101.947 Seconds
(without alignments)
53.280 Million cell updates/sec

Title: US-10-774-242A-6

Perfect score: 62

Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications_AA_Main:

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pap:
5: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pap:
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	13	4	US-10-774-242-6
2	62	100.0	760	3	US-09-265-606-2
3	62	100.0	760	4	US-10-177-293-136
4	62	100.0	760	4	US-10-301-822-55
5	62	100.0	760	5	US-10-723-860-4171
6	62	100.0	760	5	US-10-884-070A-13
7	49	79.0	136	4	US-10-767-701-53059
8	49	79.0	187	4	US-10-425-115-286978
9	49	79.0	562	4	US-10-425-114-52777
10	49	79.0	604	4	US-10-425-115-286982
11	49	79.0	633	4	US-10-437-963-186821
12	43	69.4	223	4	US-10-335-977-8596
13	43	69.4	296	4	US-10-335-977-8597
14	43	69.4	346	4	US-10-012-819-82
15	43	69.4	706	4	US-10-172-094-6
16	43	69.4	806	4	US-10-335-977-8599
17	42	67.7	17	4	US-10-419-462-9
18	42	67.7	17	5	US-10-782-968-9
19	42	67.7	18	4	US-10-419-462-10
20	42	67.7	18	5	US-10-782-968-10
21	42	67.7	98	4	US-10-419-462-42
22	42	67.7	98	5	US-10-782-968-42
23	42	67.7	384	4	US-10-424-599-257798
24	42	67.7	432	5	US-10-741-600-1020
25	42	67.7	432	5	US-10-741-600-1022
26	42	67.7	466	3	US-09-925-301-1047
27	42	67.7	535	5	US-10-723-860-4109

28	42	67.7	584	4	US-10-425-115-226623	Sequence 226623,
29	42	67.7	625	4	US-10-425-114-64494	Sequence 64494, A
30	42	67.7	831	3	US-09-939-853A-97	Sequence 97, Appl
31	42	67.7	1150	4	US-10-296-733-1	Sequence 1, Appl
32	42	67.7	1152	3	US-09-919-603-1	Sequence 1, Appl
33	42	67.7	1169	5	US-10-317-821B-7	Sequence 7, Appl
34	42	67.7	1170	4	US-10-020-141-12	Sequence 12, Appl
35	42	67.7	1170	4	US-10-017-721-2	Sequence 2, Appl
36	42	67.7	1170	4	US-10-021-660-114	Sequence 114, App
37	42	67.7	1170	4	US-10-008-093-2	Sequence 2, Appl
38	42	67.7	1170	4	US-10-295-027-1170	Sequence 1170, Ap
39	42	67.7	1170	4	US-10-211-462-38	Sequence 38, Appl
40	42	67.7	1170	4	US-10-231-956A-482	Sequence 482, App
41	42	67.7	1170	4	US-10-419-462-38	Sequence 38, Appl
42	42	67.7	1170	5	US-10-741-600-1018	Sequence 1018, Ap
43	42	67.7	1170	5	US-10-741-600-1019	Sequence 1019, Ap
44	42	67.7	1170	5	US-10-741-600-1021	Sequence 1021, Ap
45	42	67.7	1170	5	US-10-782-968-38	Sequence 38, Appl
46	42	67.7	1170	5	US-10-849-989-44	Sequence 44, Appl
47	42	67.7	1170	5	US-10-631-467-548	Sequence 548, App
48	42	67.7	1170	6	US-11-037-713-51	Sequence 51, Appl
49	41	66.1	176	4	US-10-335-977-7129	Sequence 7129, Ap
50	41	66.1	790	6	US-11-097-143-2913	Sequence 2913, Ap
51	41	66.1	847	3	US-09-815-242-11419	Sequence 11419, A
52	41	66.1	847	3	US-09-815-242-11577	Sequence 11577, A
53	41	66.1	847	3	US-09-895-913A-164	Sequence 164, App
54	41	66.1	847	4	US-10-282-122A-58924	Sequence 58924, A
55	41	66.1	847	4	US-10-335-977-7131	Sequence 7131, Ap
56	40	64.5	218	4	US-10-036-869-1	Sequence 1, Appl
57	40	64.5	245	4	US-10-205-194-99	Sequence 99, Appl
58	40	64.5	275	4	US-10-767-701-42480	Sequence 42480, A
59	40	64.5	386	4	US-10-425-115-226060	Sequence 226060,
60	40	64.5	391	3	US-09-815-242-13407	Sequence 13407, A
61	40	64.5	391	5	US-10-472-928-1398	Sequence 1398, Ap
62	40	64.5	397	4	US-10-282-122A-73879	Sequence 73879, A
63	40	64.5	397	4	US-10-474-776-276	Sequence 276, App
64	40	64.5	403	5	US-10-617-320-3191	Sequence 3191, Ap
65	40	64.5	440	4	US-10-425-114-68064	Sequence 68064, A
66	40	64.5	441	4	US-10-036-869-3	Sequence 3, Appl
67	40	64.5	486	4	US-10-425-115-227767	Sequence 227767,
68	40	64.5	517	4	US-10-425-114-56919	Sequence 56919, A
69	40	64.5	624	5	US-10-732-923-23080	Sequence 23080, A
70	40	64.5	758	4	US-10-087-192-219	Sequence 219, App
71	40	64.5	765	4	US-10-437-963-177472	Sequence 177472,
72	40	64.5	931	4	US-10-437-963-140717	Sequence 140717,
73	40	64.5	959	4	US-10-080-608A-67	Sequence 67, Appl
74	40	64.5	959	4	US-10-370-685-156	Sequence 156, App
75	40	64.5	961	4	US-10-080-608A-66	Sequence 66, Appl
76	40	64.5	961	4	US-10-370-685-155	Sequence 155, App
77	40	64.5	962	4	US-10-087-192-222	Sequence 222, App
78	40	64.5	962	5	US-10-627-311-3	Sequence 3, Appl
79	40	64.5	1279	6	US-11-097-143-24450	Sequence 24450, A
80	40	64.5	2285	3	US-09-932-183A-2	Sequence 2, Appl
81	40	64.5	2285	5	US-10-927-615-2	Sequence 2, Appl
82	40	64.5	2285	5	US-10-927-590-2	Sequence 2, Appl
83	40	64.5	2285	5	US-10-926-729-2	Sequence 2, Appl
84	39	62.9	50	3	US-09-864-761-39963	Sequence 7, Appl
85	39	62.9	161	4	US-10-792-063-7	Sequence 39963, A
86	39	62.9	233	3	US-09-815-242-10842	Sequence 10842, A
87	39	62.9	233	4	US-10-282-122A-57284	Sequence 57284, A
88	39	62.9	368	4	US-10-792-063-5	Sequence 5, Appl
89	39	62.9	405	3	US-09-895-913A-234	Sequence 234, App
90	39	62.9	405	4	US-10-335-977-7695	Sequence 7695, Ap
91	39	62.9	411	4	US-10-189-388-1	Sequence 1, Appl
92	39	62.9	411	4	US-10-408-305-1	Sequence 1, Appl
93	39	62.9	412	5	US-10-948-649-12	Sequence 12, Appl
94	39	62.9	419	3	US-09-863-049A-2	Sequence 2, Appl
95	39	62.9	419	3	US-09-851-673-2	Sequence 2, Appl
96	39	62.9	419	5	US-10-761-370-4	Sequence 4, Appl
97	39	62.9	419	5	US-10-760-678-2	Sequence 2, Appl
98	39	62.9	419	5	US-10-948-649-18	Sequence 18, Appl
99	39	62.9	420	4	US-10-425-114-60225	Sequence 60225, A
100	39	62.9	425	4	US-10-108-260A-4381	Sequence 4381, Ap

101	39	62.9	425	5	US-10-959-539-46	Sequence 46, Appl	174	37	59.7	348	5	US-10-739-930-5818	Sequence 5818, Ap
102	39	62.9	438	4	US-10-289-762-771	Sequence 771, App	175	37	59.7	381	4	US-10-425-114-43523	Sequence 43523, A
103	39	62.9	575	4	US-10-369-493-18999	Sequence 18999, A	176	37	59.7	407	4	US-10-437-963-203372	Sequence 203372, A
104	39	62.9	656	5	US-10-732-923-23044	Sequence 23044, A	177	37	59.7	416	4	US-10-282-122A-52225	Sequence 52225, A
105	39	62.9	678	4	US-10-264-049-3653	Sequence 2653, Ap	178	37	59.7	520	5	US-10-450-763-48436	Sequence 48436, A
106	39	62.9	790	4	US-10-283-132A-52378	Sequence 52378, A	179	37	59.7	530	3	US-09-861-451A-18	Sequence 18, Appl
107	39	62.9	797	6	US-11-097-143-7926	Sequence 7926, Ap	180	37	59.7	588	5	US-10-739-930-6861	Sequence 6861, Ap
108	39	62.9	867	4	US-10-282-122A-69006	Sequence 69006, Ap	181	37	59.7	654	4	US-10-425-115-251215	Sequence 251215, A
109	39	62.9	3225	4	US-10-408-765A-254	Sequence 254, App	182	37	59.7	656	4	US-10-437-963-187923	Sequence 187923, A
110	39	62.9	20	3	US-09-863-049A-76	Sequence 76, Appl	183	37	59.7	657	4	US-10-425-115-251213	Sequence 251213, A
111	38	61.3	51	4	US-10-094-944-31	Sequence 31, Appl	184	37	59.7	682	4	US-10-425-114-55123	Sequence 55125, A
112	38	61.3	96	4	US-10-424-599-192918	Sequence 192918, A	185	37	59.7	762	4	US-10-425-114-65961	Sequence 65961, A
113	38	61.3	145	4	US-10-424-599-192918	Sequence 192918, A	186	37	59.7	785	4	US-10-425-115-291758	Sequence 291758, A
114	38	61.3	154	4	US-10-437-963-134039	Sequence 134039, A	187	37	59.7	791	4	US-10-425-114-70507	Sequence 70507, A
115	38	61.3	160	4	US-10-335-977-7349	Sequence 7349, Ap	188	37	59.7	909	5	US-10-732-923-14212	Sequence 14212, A
116	38	61.3	221	5	US-10-774-355A-2107	Sequence 2107, Ap	189	37	59.7	966	5	US-10-719-993-597	Sequence 597, App
117	38	61.3	225	4	US-10-767-701-36534	Sequence 36534, A	190	37	59.7	1279	3	US-09-882-227-388	Sequence 388, App
118	38	61.3	225	4	US-10-425-115-256809	Sequence 256809, A	191	37	59.7	1847	4	US-10-369-493-1075	Sequence 1075, Ap
119	38	61.3	308	5	US-10-774-355A-1541	Sequence 1541, Ap	192	37	59.7	2954	5	US-10-650-280-1	Sequence 1, Appl
120	38	61.3	328	5	US-10-774-355A-2513	Sequence 2513, Ap	193	36	58.1	71	4	US-10-029-386-27614	Sequence 27614, A
121	38	61.3	368	4	US-10-032-585-7910	Sequence 7910, Ap	194	36	58.1	117	4	US-10-425-115-306096	Sequence 306096, A
122	38	61.3	373	4	US-10-335-977-6232	Sequence 6232, Ap	195	36	58.1	119	4	US-10-437-963-190588	Sequence 190588, A
123	38	61.3	376	4	US-10-335-977-6232	Sequence 6233, Ap	196	36	58.1	129	4	US-10-767-701-59523	Sequence 59523, A
124	38	61.3	394	4	US-10-437-963-162397	Sequence 162397, A	197	36	58.1	136	4	US-10-425-115-230753	Sequence 230753, A
125	38	61.3	430	4	US-10-369-493-2978	Sequence 2978, Ap	198	36	58.1	141	4	US-10-424-599-207564	Sequence 207564, A
126	38	61.3	448	4	US-10-282-122A-55131	Sequence 55131, A	199	36	58.1	142	4	US-10-767-701-61508	Sequence 61508, A
127	38	61.3	536	4	US-10-282-122A-68830	Sequence 68830, A	200	36	58.1	147	3	US-09-925-299-1119	Sequence 1119, Ap
128	38	61.3	545	4	US-10-424-599-192917	Sequence 192917, A	201	36	58.1	147	3	US-09-925-299-1119	Sequence 1119, Ap
129	38	61.3	609	4	US-10-369-493-1580	Sequence 1580, Ap	202	36	58.1	152	6	US-11-087-143-13014	Sequence 13014, A
130	38	61.3	638	5	US-10-732-923-18608	Sequence 18608, A	203	36	58.1	155	4	US-10-282-122A-52093	Sequence 52093, A
131	38	61.3	663	4	US-10-335-977-7351	Sequence 7351, Ap	204	36	58.1	156	4	US-10-424-599-169411	Sequence 169411, A
132	38	61.3	684	4	US-10-437-963-127649	Sequence 127649, A	205	36	58.1	159	5	US-10-732-923-14675	Sequence 14675, A
133	38	61.3	762	3	US-09-815-242-11519	Sequence 11519, A	206	36	58.1	170	4	US-10-437-963-150100	Sequence 150100, A
134	38	61.3	762	4	US-10-282-122A-58754	Sequence 58754, A	207	36	58.1	174	4	US-10-425-115-204593	Sequence 204593, A
135	38	61.3	762	4	US-10-335-977-7352	Sequence 7352, Ap	208	36	58.1	176	4	US-11-087-143-39234	Sequence 39234, A
136	38	61.3	779	5	US-10-756-149-5661	Sequence 5661, Ap	209	36	58.1	178	4	US-10-424-599-232411	Sequence 232411, A
137	38	61.3	790	4	US-10-369-493-3803	Sequence 3803, Ap	210	36	58.1	178	4	US-10-437-963-187256	Sequence 187256, A
138	38	61.3	843	5	US-10-991-321-6	Sequence 6, Appl	211	36	58.1	195	5	US-10-732-923-17809	Sequence 17809, Ap
139	38	61.3	833	5	US-10-989-891-138	Sequence 138, App	212	36	58.1	201	4	US-10-437-963-187256	Sequence 187256, A
140	38	61.3	932	3	US-09-815-242-12615	Sequence 12615, A	213	36	58.1	204	4	US-10-424-599-282842	Sequence 282842, A
141	38	61.3	932	4	US-10-301-997-45	Sequence 45, Appl	214	36	58.1	209	4	US-10-437-963-187256	Sequence 187256, A
142	38	61.3	932	4	US-10-818-509-45	Sequence 45, Appl	215	36	58.1	217	4	US-10-335-977-7818	Sequence 7818, Ap
143	38	61.3	1381	6	US-11-097-143-582	Sequence 552, App	216	36	58.1	217	5	US-10-450-763-57103	Sequence 57103, A
144	38	61.3	1527	5	US-10-732-923-3354	Sequence 3354, Ap	217	36	58.1	227	3	US-09-923-284-1	Sequence 1, Appl
145	38	61.3	1731	4	US-10-437-963-193200	Sequence 193200, A	218	36	58.1	245	4	US-10-080-608A-65	Sequence 65, Appl
146	37	59.7	58	4	US-10-282-122A-52591	Sequence 52591, A	219	36	58.1	245	4	US-10-370-685-154	Sequence 154, Appl
147	37	59.7	72	5	US-10-856-499-621	Sequence 621, App	220	36	58.1	245	5	US-10-717-665-34	Sequence 34, Appl
148	37	59.7	90	4	US-10-424-599-231423	Sequence 231423, A	221	36	58.1	281	4	US-10-080-608A-63	Sequence 63, Appl
149	37	59.7	92	4	US-10-424-599-260755	Sequence 260755, A	222	36	58.1	281	4	US-10-370-685-152	Sequence 152, App
150	37	59.7	112	4	US-10-424-599-232843	Sequence 232843, A	223	36	58.1	281	4	US-10-437-963-199294	Sequence 199294, A
151	37	59.7	128	4	US-10-425-115-332209	Sequence 332209, A	224	36	58.1	292	4	US-10-437-963-187257	Sequence 187257, A
152	37	59.7	153	4	US-10-425-115-332207	Sequence 332207, A	225	36	58.1	294	4	US-10-774-355A-1533	Sequence 1533, Ap
153	37	59.7	185	4	US-10-425-114-39974	Sequence 39974, A	226	36	58.1	305	5	US-10-774-355A-1539	Sequence 1539, Ap
154	37	59.7	210	4	US-10-425-115-332304	Sequence 332304, A	227	36	58.1	306	4	US-10-106-698-4835	Sequence 4835, Ap
155	37	59.7	210	4	US-10-425-115-332205	Sequence 332205, A	228	36	58.1	306	5	US-10-473-127-407	Sequence 407, App
156	37	59.7	222	4	US-10-425-114-65746	Sequence 65746, A	229	36	58.1	307	5	US-10-774-355A-2100	Sequence 2100, Ap
157	37	59.7	222	4	US-10-424-599-271480	Sequence 271480, A	230	36	58.1	311	5	US-10-774-355A-2308	Sequence 2308, Ap
158	37	59.7	231	4	US-10-369-493-3217	Sequence 3217, Ap	231	36	58.1	311	5	US-10-774-355A-1537	Sequence 1537, Ap
159	37	59.7	233	4	US-10-425-114-67549	Sequence 67549, A	232	36	58.1	312	5	US-10-774-355A-1539	Sequence 1539, Ap
160	37	59.7	251	4	US-10-425-114-57068	Sequence 57068, A	233	36	58.1	312	5	US-10-774-355A-1540	Sequence 1540, Ap
161	37	59.7	255	4	US-10-425-114-66645	Sequence 66645, A	234	36	58.1	312	5	US-10-774-355A-2031	Sequence 2031, Ap
162	37	59.7	271	4	US-10-767-701-40429	Sequence 40429, A	235	36	58.1	331	4	US-10-292-798-496	Sequence 496, App
163	37	59.7	273	4	US-10-767-701-43113	Sequence 43113, A	236	36	58.1	331	4	US-10-467-252-27	Sequence 27, Appl
164	37	59.7	289	5	US-10-719-993-643	Sequence 643, App	237	36	58.1	334	4	US-10-425-114-56387	Sequence 56387, A
165	37	59.7	294	4	US-10-278-173-86	Sequence 86, Appl	238	36	58.1	334	4	US-10-425-115-368875	Sequence 368875, A
166	37	59.7	294	4	US-10-295-403-126	Sequence 126, App	239	36	58.1	343	4	US-10-085-198-206	Sequence 206, App
167	37	59.7	294	4	US-10-412-698B-470	Sequence 470, App	240	36	58.1	357	4	US-10-369-493-5418	Sequence 5418, Ap
168	37	59.7	301	4	US-10-342-224-74	Sequence 74, Appl	241	36	58.1	357	4	US-10-017-161-482	Sequence 482, App
169	37	59.7	305	5	US-10-774-355A-2133	Sequence 2133, Ap	242	36	58.1	392	4	US-10-437-963-165680	Sequence 165680, A
170	37	59.7	310	4	US-10-425-114-38523	Sequence 38523, A	243	36	58.1	400	4	US-10-425-115-342645	Sequence 342645, A
171	37	59.7	310	5	US-10-732-923-13908	Sequence 13908, A	244	36	58.1	412	4	US-10-425-114-43254	Sequence 43254, A
172	37	59.7	314	5	US-10-774-355A-2197	Sequence 2197, Ap	245	36	58.1	412	4	US-10-425-114-68952	Sequence 68952, A
173	37	59.7	319	5	US-10-774-355A-1851	Sequence 1851, Ap	246	36	58.1	413	4	US-10-425-114-37572	Sequence 37572, A

247	36	58.1	419	5	US-10-450-763-52471	Sequence 52471, A	320	35	56.5	112	3	US-09-833-245-727	Sequence 727, App
248	36	58.1	420	4	US-10-369-493-1771	Sequence 1771, App	321	35	56.5	112	4	US-10-437-963-137694	Sequence 137694, A
249	36	58.1	435	3	US-09-895-913A-146	Sequence 146, App	322	35	56.5	118	4	US-10-767-701-32643	Sequence 32643, A
250	36	58.1	435	4	US-10-335-977-7949	Sequence 7949, App	323	35	56.5	122	4	US-10-112-944-430	Sequence 430, App
251	36	58.1	454	4	US-10-425-114-63299	Sequence 63299, A	324	35	56.5	123	3	US-09-833-245-1228	Sequence 1228, App
252	36	58.1	474	4	US-10-282-122A-70588	Sequence 70588, A	325	35	56.5	123	3	US-09-833-245-1229	Sequence 1229, App
253	36	58.1	478	4	US-10-724-972A-6580	Sequence 6580, App	326	35	56.5	134	4	US-10-767-701-41505	Sequence 41505, A
254	36	58.1	485	4	US-10-282-122A-52735	Sequence 52735, A	327	35	56.5	139	3	US-09-832-129-53	Sequence 53, Appl
255	36	58.1	491	4	US-10-179-046-2	Sequence 2, Appl	328	35	56.5	139	3	US-09-833-245-726	Sequence 726, App
256	36	58.1	496	4	US-10-425-114-71015	Sequence 71015, A	329	35	56.5	139	3	US-09-833-245-2247	Sequence 2247, App
257	36	58.1	509	4	US-10-437-963-156894	Sequence 156894, A	330	35	56.5	139	4	US-10-733-368-53	Sequence 53, Appl
258	36	58.1	520	4	US-10-425-115-332309	Sequence 332309, A	331	35	56.5	151	4	US-10-425-115-291844	Sequence 291844, A
259	36	58.1	524	6	US-11-097-143-4239	Sequence 4239, App	332	35	56.5	153	4	US-10-282-122A-68730	Sequence 68730, A
260	36	58.1	547	4	US-10-425-114-64788	Sequence 64788, A	333	35	56.5	189	4	US-11-097-143-35504	Sequence 35504, A
261	36	58.1	560	4	US-10-435-696-44	Sequence 44, Appl	334	35	56.5	189	4	US-10-425-115-264511	Sequence 264511, A
262	36	58.1	560	4	US-10-408-765A-393	Sequence 393, App	335	35	56.5	195	4	US-10-335-977-8101	Sequence 8101, App
263	36	58.1	560	4	US-10-693-999-6	Sequence 6, Appl	336	35	56.5	197	4	US-10-335-977-8101	Sequence 8101, App
264	36	58.1	560	4	US-10-693-999-7	Sequence 7, Appl	337	35	56.5	202	4	US-10-017-161-1478	Sequence 1478, App
265	36	58.1	560	5	US-10-473-127-405	Sequence 405, App	338	35	56.5	202	4	US-10-292-798-1186	Sequence 1186, App
266	36	58.1	560	5	US-10-473-127-406	Sequence 406, App	339	35	56.5	209	4	US-10-424-599-265686	Sequence 265686, A
267	36	58.1	560	5	US-10-473-127-408	Sequence 408, App	340	35	56.5	210	4	US-10-437-963-191664	Sequence 191664, A
268	36	58.1	560	5	US-10-473-127-409	Sequence 409, App	341	35	56.5	211	4	US-10-425-114-67965	Sequence 67965, A
269	36	58.1	560	5	US-10-473-127-410	Sequence 410, App	342	35	56.5	215	4	US-10-767-701-51911	Sequence 51911, A
270	36	58.1	560	5	US-10-473-127-411	Sequence 411, App	343	35	56.5	215	4	US-10-767-701-32376	Sequence 32376, A
271	36	58.1	572	4	US-10-425-115-340590	Sequence 340590, A	344	35	56.5	216	4	US-10-425-115-190769	Sequence 190769, A
272	36	58.1	575	4	US-10-724-972A-7483	Sequence 7483, App	345	35	56.5	220	4	US-10-425-114-64931	Sequence 64931, A
273	36	58.1	582	4	US-10-282-122A-52025	Sequence 52025, A	346	35	56.5	225	4	US-10-424-599-200502	Sequence 200502, A
274	36	58.1	582	4	US-10-411-910A-23	Sequence 23, Appl	347	35	56.5	234	5	US-10-733-923-15843	Sequence 15843, A
275	36	58.1	591	4	US-10-425-114-48621	Sequence 48621, A	348	35	56.5	235	4	US-10-282-122A-57689	Sequence 57689, A
276	36	58.1	592	4	US-10-425-115-340601	Sequence 340601, A	349	35	56.5	239	6	US-11-097-143-37722	Sequence 37722, A
277	36	58.1	602	5	US-10-732-923-13440	Sequence 13440, A	350	35	56.5	247	4	US-10-369-493-11200	Sequence 11200, A
278	36	58.1	613	3	US-09-971-309-5	Sequence 5, Appl	351	35	56.5	247	4	US-10-781-014-784	Sequence 784, App
279	36	58.1	622	6	US-11-013-314-40	Sequence 40, Appl	352	35	56.5	252	4	US-10-425-115-303049	Sequence 303049, A
280	36	58.1	622	6	US-11-097-143-15240	Sequence 15240, A	353	35	56.5	255	4	US-10-094-944-29	Sequence 29, Appl
281	36	58.1	649	4	US-10-282-122A-71864	Sequence 71864, A	354	35	56.5	257	4	US-10-425-115-274780	Sequence 274780, A
282	36	58.1	685	4	US-10-369-493-2526	Sequence 2526, App	355	35	56.5	277	3	US-09-934-455-246	Sequence 246, App
283	36	58.1	714	4	US-10-335-977-6468	Sequence 6468, App	356	35	56.5	277	4	US-10-286-264-50	Sequence 50, Appl
284	36	58.1	716	4	US-10-437-963-140947	Sequence 140947, A	357	35	56.5	277	4	US-10-225-066A-1000	Sequence 1000, App
285	36	58.1	757	4	US-10-425-115-226090	Sequence 226090, A	358	35	56.5	277	4	US-10-374-780A-306	Sequence 306, App
286	36	58.1	761	4	US-10-369-493-129	Sequence 129, App	359	35	56.5	277	5	US-10-413-699B-928	Sequence 928, App
287	36	58.1	784	4	US-10-041-018-383	Sequence 383, App	360	35	56.5	288	4	US-10-225-066A-1000	Sequence 1000, App
288	36	58.1	784	4	US-10-041-018-384	Sequence 384, App	361	35	56.5	288	4	US-10-282-122A-46286	Sequence 46286, A
289	36	58.1	790	4	US-10-225-066A-810	Sequence 810, App	362	35	56.5	299	4	US-10-425-114-55024	Sequence 55024, A
290	36	58.1	790	4	US-10-374-780A-2384	Sequence 2384, App	363	35	56.5	305	3	US-09-906-179A-215	Sequence 215, App
291	36	58.1	790	5	US-10-225-066A-810	Sequence 810, App	364	35	56.5	305	4	US-10-671-403-126	Sequence 126, App
292	36	58.1	790	5	US-10-396-058-42	Sequence 42, Appl	365	35	56.5	305	4	US-10-671-403-126	Sequence 126, App
293	36	58.1	812	4	US-10-437-963-161917	Sequence 161917, A	366	35	56.5	305	4	US-10-670-844-126	Sequence 126, App
294	36	58.1	831	3	US-09-939-853A-98	Sequence 98, Appl	367	35	56.5	305	4	US-10-671-134-126	Sequence 126, App
295	36	58.1	849	4	US-10-425-115-273371	Sequence 273371, A	368	35	56.5	305	4	US-10-673-098-126	Sequence 126, App
296	36	58.1	856	4	US-10-099-322-116	Sequence 116, App	369	35	56.5	305	4	US-10-673-127-126	Sequence 126, App
297	36	58.1	856	4	US-10-044-564-116	Sequence 4146, App	370	35	56.5	305	4	US-10-670-817-126	Sequence 126, App
298	36	58.1	856	6	US-11-097-143-4146	Sequence 4146, App	371	35	56.5	305	4	US-10-673-119-126	Sequence 126, App
299	36	58.1	857	4	US-10-425-114-60120	Sequence 60120, A	372	35	56.5	305	4	US-10-746-167-89	Sequence 89, Appl
300	36	58.1	954	4	US-10-437-963-156892	Sequence 156892, A	373	35	56.5	305	5	US-10-671-207-126	Sequence 126, App
301	36	58.1	978	5	US-10-732-923-3309	Sequence 3309, App	374	35	56.5	305	5	US-10-673-120-126	Sequence 126, App
302	36	58.1	1050	5	US-10-756-149-5594	Sequence 5594, App	375	35	56.5	305	5	US-10-671-413-126	Sequence 126, App
303	36	58.1	1066	5	US-10-450-763-46229	Sequence 46229, A	376	35	56.5	305	5	US-10-671-859-126	Sequence 126, App
304	36	58.1	1071	4	US-10-437-963-156895	Sequence 156895, A	377	35	56.5	305	5	US-10-671-106-126	Sequence 126, App
305	36	58.1	1170	5	US-10-631-467-1376	Sequence 1376, App	378	35	56.5	310	4	US-10-282-122A-72138	Sequence 72138, A
306	36	58.1	1278	5	US-10-828-924-79	Sequence 79, Appl	379	35	56.5	311	4	US-10-156-761-10910	Sequence 10910, A
307	36	58.1	1289	4	US-10-424-599-204412	Sequence 204412, A	380	35	56.5	316	5	US-10-774-355A-2249	Sequence 2249, App
308	36	58.1	1550	4	US-10-437-963-112048	Sequence 112048, A	381	35	56.5	317	4	US-10-437-963-187215	Sequence 187215, A
309	36	58.1	1738	5	US-10-732-923-8335	Sequence 8335, App	382	35	56.5	332	4	US-10-282-122A-59407	Sequence 59407, A
310	36	58.1	1982	5	US-10-450-763-46763	Sequence 46763, A	383	35	56.5	334	4	US-10-424-599-260259	Sequence 260259, A
311	36	58.1	1983	5	US-10-450-763-50015	Sequence 50015, A	384	35	56.5	336	3	US-09-789-919-52	Sequence 52, Appl
312	36	58.1	1983	5	US-10-723-860-1890	Sequence 1890, App	385	35	56.5	336	3	US-10-282-122A-61079	Sequence 61079, A
313	36	58.1	3321	6	US-11-075-234-4	Sequence 4, Appl	386	35	56.5	336	3	US-10-732-923-7790	Sequence 7790, App
314	36	58.1	473	4	US-10-369-493-13571	Sequence 13571, A	387	35	56.5	353	5	US-10-732-923-7965	Sequence 7965, App
315	35	57.3	50	3	US-09-864-761-45347	Sequence 45347, A	388	35	56.5	353	5	US-10-732-923-7965	Sequence 7965, App
316	35	56.5	55	3	US-09-864-761-35539	Sequence 35539, A	389	35	56.5	362	4	US-10-282-122A-53180	Sequence 53180, A
317	35	56.5	79	4	US-10-767-701-54559	Sequence 54559, A	390	35	56.5	363	3	US-09-832-129-34	Sequence 34, Appl
318	35	56.5	109	4	US-10-425-115-263571	Sequence 263571, A	391	35	56.5	363	3	US-09-833-245-2248	Sequence 2248, App
319	35	56.5	112	3	US-09-833-245-723	Sequence 723, App	392	35	56.5				

393	35	56.5	363	4	US-10-733-368-34	Sequence 34, Appl	466	35	56.5	1024	4	US-10-177-293-252	Sequence 252, App
394	35	56.5	369	5	US-10-470-088B-561	Sequence 561, App	467	35	56.5	1027	4	US-10-335-977-8086	Sequence 8086, Ap
395	35	56.5	370	4	US-10-724-972A-4799	Sequence 4799, Ap	468	35	56.5	1030	4	US-10-425-114-62748	Sequence 62748, A
396	35	56.5	371	5	US-10-472-928-2246	Sequence 2246, Ap	469	35	56.5	1080	4	US-10-425-115-231230	Sequence 231230, A
397	35	56.5	375	3	US-09-815-242-13314	Sequence 13314, A	470	35	56.5	1104	5	US-10-473-127-1367	Sequence 1367, Ap
398	35	56.5	375	4	US-10-424-599-328292	Sequence 328292, A	471	35	56.5	1104	5	US-10-473-127-1369	Sequence 1369, Ap
399	35	56.5	376	4	US-10-282-122A-76580	Sequence 76580, A	472	35	56.5	1104	5	US-10-473-127-1370	Sequence 1370, Ap
400	35	56.5	376	4	US-10-282-122A-76580	Sequence 76580, A	473	35	56.5	1104	5	US-10-723-860-2274	Sequence 2274, Ap
401	35	56.5	376	4	US-10-335-977-9342	Sequence 9342, Ap	474	35	56.5	1104	5	US-10-756-149-5320	Sequence 5320, Ap
402	35	56.5	378	5	US-10-617-320-4822	Sequence 4822, Ap	475	35	56.5	1122	5	US-10-771-708-3	Sequence 3, Appli
403	35	56.5	382	3	US-09-881-752A-224	Sequence 224, App	476	35	56.5	1131	6	US-11-097-143-40503	Sequence 40503, A
404	35	56.5	384	4	US-10-335-977-9343	Sequence 9343, Ap	477	35	56.5	1154	4	US-10-369-493-5385	Sequence 5385, Ap
405	35	56.5	395	5	US-10-739-930-8158	Sequence 8158, Ap	478	35	56.5	1177	5	US-10-732-923-3293	Sequence 3293, Ap
406	35	56.5	403	4	US-10-424-599-200504	Sequence 200504, A	479	35	56.5	1209	6	US-11-097-143-4206	Sequence 4206, Ap
407	35	56.5	409	6	US-11-097-143-15057	Sequence 15057, A	480	35	56.5	1251	4	US-10-452-024-127	Sequence 127, App
408	35	56.5	424	4	US-10-424-599-256696	Sequence 256696, A	481	35	56.5	1254	4	US-10-369-493-22131	Sequence 22131, A
409	35	56.5	431	6	US-11-097-143-41865	Sequence 41865, A	482	35	56.5	1261	4	US-10-177-980-2	Sequence 2, Appli
410	35	56.5	435	4	US-10-369-493-141	Sequence 141, App	483	35	56.5	1261	4	US-10-072-012-733	Sequence 733, App
411	35	56.5	446	4	US-10-282-122A-71116	Sequence 71116, A	484	35	56.5	1261	4	US-10-648-593-192	Sequence 12754, A
412	35	56.5	450	4	US-10-437-963-152372	Sequence 152372, A	485	35	56.5	1268	5	US-10-732-923-12754	Sequence 1217, Ap
413	35	56.5	452	4	US-10-724-972A-4374	Sequence 4374, Ap	486	35	56.5	1275	4	US-10-408-765A-2117	Sequence 1215, Ap
414	35	56.5	455	4	US-10-424-599-221866	Sequence 221866, A	487	35	56.5	1327	4	US-10-408-765A-2117	Sequence 1215, Ap
415	35	56.5	460	4	US-10-424-599-260258	Sequence 260258, A	488	35	56.5	1345	4	US-10-437-963-160952	Sequence 160952, A
416	35	56.5	464	4	US-10-369-493-23304	Sequence 23304, A	489	35	56.5	1396	6	US-11-097-143-20388	Sequence 20388, A
417	35	56.5	465	4	US-10-732-923-23683	Sequence 23683, A	490	35	56.5	1434	4	US-10-282-122A-76740	Sequence 76740, A
418	35	56.5	465	4	US-10-108-260A-4565	Sequence 4565, Ap	491	35	56.5	1557	4	US-10-369-493-2224	Sequence 2224, Ap
419	35	56.5	466	3	US-09-826-191-9	Sequence 9, Appli	492	35	56.5	1786	3	US-09-742-096-3	Sequence 3, Appli
420	35	56.5	466	4	US-10-263-666-9	Sequence 9, Appli	493	35	56.5	1787	4	US-10-415-253-2	Sequence 2, Appli
421	35	56.5	466	4	US-10-261-481-9	Sequence 9, Appli	494	35	56.5	1809	5	US-10-745-237-54	Sequence 54, Appl
422	35	56.5	481	3	US-09-901-884-7	Sequence 7, Appli	495	35	56.5	1809	6	US-11-097-143-13683	Sequence 13683, A
423	35	56.5	481	5	US-10-694-779-7	Sequence 7, Appli	496	35	56.5	2025	4	US-10-788-792-180	Sequence 180, App
424	35	56.5	483	3	US-09-738-626-4842	Sequence 4842, Ap	497	35	56.5	2258	3	US-09-892-635A-25	Sequence 25, Appli
425	35	56.5	483	4	US-10-425-114-39298	Sequence 39298, A	498	35	56.5	2325	4	US-10-663-433-2	Sequence 2, Appli
426	35	56.5	483	4	US-10-781-014-762	Sequence 762, App	499	35	56.5	2384	4	US-10-335-977-8087	Sequence 8087, Ap
427	35	56.5	493	4	US-10-425-114-56983	Sequence 56983, A	500	35	56.5	2440	4	US-10-335-977-8088	Sequence 8088, Ap
428	35	56.5	504	4	US-10-072-012-863	Sequence 863, App							
429	35	56.5	521	4	US-10-282-122A-53435	Sequence 53435, A							
430	35	56.5	521	4	US-10-424-599-242324	Sequence 242324, A							
431	35	56.5	569	6	US-11-097-143-36531	Sequence 36531, A							
432	35	56.5	587	4	US-10-424-599-202287	Sequence 202287, A							
433	35	56.5	610	5	US-10-732-923-18636	Sequence 18636, A							
434	35	56.5	613	4	US-10-425-115-222182	Sequence 222182, A							
435	35	56.5	625	4	US-10-437-963-144600	Sequence 144600, A							
436	35	56.5	647	4	US-10-369-493-6272	Sequence 6272, Ap							
437	35	56.5	651	4	US-10-425-115-356387	Sequence 356387, A							
438	35	56.5	654	4	US-10-425-114-65105	Sequence 65105, A							
439	35	56.5	658	4	US-10-425-114-54318	Sequence 54318, A							
440	35	56.5	679	4	US-10-228-063-47	Sequence 47, Appl							
441	35	56.5	703	4	US-10-425-115-253968	Sequence 253968, A							
442	35	56.5	714	3	US-09-881-752A-322	Sequence 322, App							
443	35	56.5	714	4	US-10-282-122A-58644	Sequence 58644, A							
444	35	56.5	714	4	US-10-282-122A-58644	Sequence 58644, A							
445	35	56.5	799	4	US-10-439-2	Sequence 2, Appli							
446	35	56.5	799	4	US-10-788-266-2	Sequence 2, Appli							
447	35	56.5	807	3	US-09-847-046-2	Sequence 4, Appli							
448	35	56.5	811	4	US-10-213-439-4	Sequence 4, Appli							
449	35	56.5	811	4	US-10-788-266-4	Sequence 4, Appli							
450	35	56.5	825	4	US-10-295-027-546	Sequence 546, App							
451	35	56.5	858	4	US-10-367-978-35	Sequence 35, Appl							
452	35	56.5	858	5	US-10-723-860-4352	Sequence 4352, Ap							
453	35	56.5	874	4	US-10-282-122A-67951	Sequence 67951, A							
454	35	56.5	880	4	US-10-369-493-21643	Sequence 21643, A							
455	35	56.5	880	5	US-10-732-923-3314	Sequence 3314, Ap							
456	35	56.5	893	4	US-10-425-115-366336	Sequence 366336, A							
457	35	56.5	913	3	US-09-801-574-65	Sequence 65, Appl							
458	35	56.5	913	4	US-10-092-900A-340	Sequence 340, App							
459	35	56.5	919	4	US-10-369-493-22825	Sequence 22825, A							
460	35	56.5	962	3	US-09-738-626-4322	Sequence 4322, Ap							
461	35	56.5	962	5	US-10-494-541-96	Sequence 96, Appl							
462	35	56.5	965	4	US-10-437-963-141471	Sequence 141471, A							
463	35	56.5	981	4	US-10-437-963-149503	Sequence 149503, A							
464	35	56.5	992	5	US-10-739-930-7903	Sequence 7903, Ap							
465	35	56.5	1001	4	US-10-425-115-231231	Sequence 231231, A							
			1014	4	US-10-437-963-156793	Sequence 156793, A							

ALIGNMENTS

RESULT 1

US-10-774-242-6

Sequence 6, Application US/10774242

Publication No. US20040203102A1

GENERAL INFORMATION:

APPLICANT: McKee, Patrick A

APPLICANT: Lee, Kyung N.

APPLICANT: Jackson, Kenneth W.

APPLICANT: Christiansen, Victoria J.

TITLE OF INVENTION: ANTIPLASMIN CLEAVING ENZYME

FILE REFERENCE: 5820.646

CURRENT APPLICATION NUMBER: US/10774.242

CURRENT FILING DATE: 2004-02-06

PRIOR APPLICATION NUMBER: 60/445,774

PRIOR FILING DATE: 2003-02-07

NUMBER OF SEQ ID NOS: 11

SOFTWARE: Patent in version 3.1

SEQ ID NO 6

LENGTH: 13

TYPE: PRT

ORGANISM: Homo sapiens

US-10-774-242-6

Query Match

Best Local Similarity

Matches

13;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

Qy

1

I

L

E

E

N

K

E

L

E

N

A

L

K

13

Db

1

I

L

E

E

N

K

E

L

E

N

A

L

K

13

RESULT 2

```

US-09-265-606-2
; Sequence 2, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020034789A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-265-606-2
Query Match 100.0%; Score 62; DB 3; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 487 ILEENKELENALK 499

RESULT 3
US-10-177-293-136
; Sequence 136, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos

US-09-265-606-2
Query Match 100.0%; Score 62; DB 3; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 487 ILEENKELENALK 499

RESULT 3
US-10-177-293-136
; Sequence 136, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos

```

```

; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-177-293-136
Query Match 100.0%; Score 62; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
Db 487 ILEENKELENALK 499

RESULT 4
US-10-301-822-55
; Sequence 55, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
;
US-10-301-822-55
Query Match 100.0%; Score 62; DB 4; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ILEENKELENALK 13
|||||
Db 487 ILEENKELENALK 499

RESULT 5

US-10-723-860-4171
; Sequence 4171, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4171
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4171

Query Match 100.0%; Score 62; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||||
Db 487 ILEENKELENALK 499

RESULT 6

US-10-884-070A-13
; Sequence 13, Application US/10884070A
; Publication No. US20050170368A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Cairns, Belinda
; APPLICANT: Dowd, Patrick
; APPLICANT: Frantz, Gretchen
; APPLICANT: French, Dorothy
; APPLICANT: Gonzalez, Lino
; APPLICANT: Polakis, Paul
; APPLICANT: Smith, Victoria
; APPLICANT: Wolf, Beni B.
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5036R1-US
; CURRENT APPLICATION NUMBER: US/10/884,070A
; PRIOR FILING DATE: 2004-07-02
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 13
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-884-070A-13

Query Match 100.0%; Score 62; DB 5; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13

Db 487 ILEENKELENALK 499
|||||

RESULT 7

US-10-767-701-53059
; Sequence 53059, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 53059
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 13239136.pcp
US-10-767-701-53059

Query Match 79.0%; Score 49; DB 4; Length 136;
Best Local Similarity 69.2%; Pred. No. 4.8;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:||||:||||:
Db 77 VLEENQSLLENALR 89

RESULT 8

US-10-425-115-286978
; Sequence 286978, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 286978
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_24818C.1.pcp
US-10-425-115-286978

Query Match 79.0%; Score 49; DB 4; Length 187;
Best Local Similarity 69.2%; Pred. No. 6.8;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:||||:||||:
Db 77 VLEENQSLLENALR 89

RESULT 9

US-10-425-114-52777
; Sequence 52777, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 52777
 LENGTH: 562
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: 700236753_FLI.pep
 US-10-425-114-52777

Query Match 79.0%; Score 49; DB 4; Length 562;
 Best Local Similarity 69.2%; Pred. No. 23;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 :||||:||||:
 DB 1 VLEENQSLENALR 13

RESULT 10
 US-10-425-115-286982
 Sequence 286982, Application US/10425115
 Publication No. US20040214272A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(53222)B
 CURRENT APPLICATION NUMBER: US/10/425,115
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 369326
 SEQ ID NO 286982
 LENGTH: 604
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: MRT4577_24821C.1.pep
 US-10-425-115-286982

Query Match 79.0%; Score 49; DB 4; Length 604;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 :||||:||||:
 DB 43 VLEENQSLENALR 55

RESULT 11
 US-10-437-963-186821
 Sequence 186821, Application US/10437963
 Publication No. US20040123343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 186821
 LENGTH: 653
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_83584C.1.pep
 US-10-437-963-186821

Query Match 79.0%; Score 49; DB 4; Length 653;
 Best Local Similarity 69.2%; Pred. No. 27;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
 :||||:||||:
 DB 83 VLEENQSLENALR 95

RESULT 12
 US-10-335-977-8596
 Sequence 8596, Application US/10335977
 Publication No. US20040052799A1
 GENERAL INFORMATION:
 APPLICANT: DOUGLAS SMITH et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 RELATING TO HELICOBACTER PYLORI FOR
 DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 10031
 CORRESPONDENCE ADDRESS:
 ADDRESSER: LAHIVE & COCKFIELD
 STREET: 28 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: Windows NT 4.0
 SOFTWARE: UNIX
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/335,977
 FILING DATE: 30-Dec-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/993,002
 FILING DATE: 17-DEC-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragoras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: GTN-018
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)742-4214
 INFORMATION FOR SEQ ID NO: 8596:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Helicobacter pylori
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1...223
 SEQUENCE DESCRIPTION: SEQ ID NO: 8596:
 US-10-335-977-8596

Query Match 69.4%; Score 43; DB 4; Length 223;
 Best Local Similarity 69.2%; Pred. No. 74;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
 |||:|:|:|:|
 Db 208 ILTESKEIEALK 220

RESULT 13
 US-10-335-977-8597
 ; Sequence 8597, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragouras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)742-4214
 ; INFORMATION FOR SEQ ID NO: 8597:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 296 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (8) LOCATION 1...296
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8597:
 US-10-335-977-8597

Query Match 69.4%; Score 43; DB 4; Length 296;
 Best Local Similarity 69.2%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
 |||:|:|:|:|
 Db 176 ILTESKEIEALK 188

RESULT 14
 US-10-012-819-82
 ; Sequence 82, Application US/10012819
 ; Publication No. US20030017478A1
 ; GENERAL INFORMATION:

; APPLICANT: Legrain, Pierre
 ; APPLICANT: Selig, Luc
 ; APPLICANT: Rain, Jean-Christophe
 ; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
 ; TITLE OF INVENTION: Pylori Protein-Protein Interactions and Applications thereof
 ; FILE REFERENCE: B5053
 ; CURRENT APPLICATION NUMBER: US/10/012,819
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: EP 99401066.8
 ; PRIOR FILING DATE: 1999-04-30
 ; NUMBER OF SEQ ID NOS: 278
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 82
 ; LENGTH: 346
 ; TYPE: PRT
 ; ORGANISM: Helicobacter pylori
 ; US-10-012-819-82

Query Match 69.4%; Score 43; DB 4; Length 346;
 Best Local Similarity 69.2%; Pred. No. 1.2e+02;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
 |||:|:|:|:|
 Db 157 ILTESKEIEALK 169

RESULT 15
 US-10-172-094-6
 ; Sequence 6, Application US/10172094
 ; Publication No. US20030161830A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JACKSON, DONALD
 ; APPLICANT: LORENZI, MATTHEW
 ; APPLICANT: ATTAR, RICARDO
 ; APPLICANT: GOTTARDIS, MARCO
 ; TITLE OF INVENTION: NOVEL HUMAN HISTONE PEPTIDES
 ; FILE REFERENCE: 3053-4145US1
 ; CURRENT APPLICATION NUMBER: US/10/172,094
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/298,296
 ; PRIOR FILING DATE: 2001-06-14
 ; NUMBER OF SEQ ID NOS: 127
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 706
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; US-10-172-094-6

Query Match 69.4%; Score 43; DB 4; Length 706;
 Best Local Similarity 81.8%; Pred. No. 2.7e+02;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LEENKELENAL 12
 |||||:|:|:|
 Db 22 LEENKELENSL 32

RESULT 16
 US-10-335-977-8599
 ; Sequence 8599, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston

STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 8599:
SEQUENCE CHARACTERISTICS:
LENGTH: 806 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...806
SEQUENCE DESCRIPTION: SEQ ID NO: 8599:
US-10-335-977-8599

Query Match 69.4%; Score 43; DB 4; Length 806;
Best Local Similarity 69.2%; Pred. No. 3.1e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
|||:|||||

DB 155 ILTSEKEIEALK 167

RESULT 17
US-10-419-462-9
; Sequence 9, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region
US-10-419-462-9

Query Match 67.7%; Score 42; DB 4; Length 17;
Best Local Similarity 61.5%; Pred. No. 6.1;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:|||||||

DB 3 VTEENKELANELR 15

RESULT 18

US-10-782-968-9
; Sequence 9, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region
US-10-782-968-9

Query Match 67.7%; Score 42; DB 5; Length 17;
Best Local Similarity 61.5%; Pred. No. 6.1;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:|||||||

DB 3 VTEENKELANELR 15

RESULT 19

US-10-419-462-10
; Sequence 10, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region which includes an N-terminal Cys added to
US-10-419-462-10

Query Match 67.7%; Score 42; DB 4; Length 18;
Best Local Similarity 61.5%; Pred. No. 6.5;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
:|||||||

DB 4 VTEENKELANELR 16

RESULT 20

US-10-782-968-10
; Sequence 10, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:

```
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region which includes an N-terminal Cys added to
; OTHER INFORMATION: aid conjugation
US-10-782-968-10

Query Match      67.7%; Score 42; DB 5; Length 18;
Best Local Similarity 61.5%; Pred. No. 6.5;
Matches      8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      4 VTEENKELANELR 16

RESULT 21
US-10-419-462-42
; Sequence 42, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus Procollagen homology domain
US-10-419-462-42

Query Match      67.7%; Score 42; DB 4; Length 98;
Best Local Similarity 61.5%; Pred. No. 43;
Matches      8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      21 VTEENKELANELR 33

RESULT 22
US-10-782-968-42
; Sequence 42, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
```

```
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thrombospondin Region plus Procollagen homology domain
US-10-782-968-42
```

```
Query Match      67.7%; Score 42; DB 5; Length 98;
Best Local Similarity 61.5%; Pred. No. 43;
Matches      8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      1 ILEENKELENALK 13
Db      21 VTEENKELANELR 33
```

RESULT 23

```
US-10-424-599-257798
; Sequence 257798, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 257798
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(384)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74815C.1.pep
US-10-424-599-257798
```

```
Query Match      67.7%; Score 42; DB 4; Length 384;
Best Local Similarity 72.7%; Pred. No. 2e+02;
Matches      8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 LEENKELENAL 12
Db      147 MEENKELESL 157
```

RESULT 24

```
US-10-741-600-1020
; Sequence 1020, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1020
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(432)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1020

Query Match 67.7%; Score 42; DB 5; Length 432;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANELR 313

RESULT 25
US-10-741-600-1022
Sequence 1022, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1022
LENGTH: 432
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(432)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-741-600-1022

Query Match 67.7%; Score 42; DB 5; Length 432;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANELR 313

RESULT 26
US-09-925-301-1047
Sequence 1047, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: ROSEN ET AL.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1047
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-301-1047

Query Match 67.7%; Score 42; DB 3; Length 466;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 354 VTEENKELANELR 366

RESULT 27
US-10-723-860-4109
Sequence 4109, Application US/10723860
Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Gineburg, Wendy M.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
FILE REFERENCE: 05882.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723,860
CURRENT FILING DATE: 2003-11-26
PRIOR APPLICATION NUMBER: 60/429,739
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4109
LENGTH: 535
TYPE: PRT
ORGANISM: Homo sapiens
US-10-723-860-4109

Query Match 67.7%; Score 42; DB 5; Length 535;
Best Local Similarity 61.5%; Pred. No. 2.8e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 108 ILEENKELEGSLQ 120

RESULT 28
US-10-425-115-226623
Sequence 226623, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 226623
LENGTH: 584
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_138267C.1.pcp
US-10-425-115-226623

Query Match 67.7%; Score 42; DB 4; Length 584;
Best Local Similarity 61.5%; Pred. No. 3.1e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 391 VLEEKAELENRMK 403

RESULT 29
US-10-425-114-64494
Sequence 64494, Application US/10425114
Publication No. US20040034888A1


```
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 64494
/ LENGTH: 625
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3060-111-G1_FLI.pep
US-10-425-114-64494

Query Match      67.7%; Score 42; DB 4; Length 625;
Best Local Similarity 61.5%; Pred. No. 3.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      432 VLEEKAELENRMK 444

RESULT 30
US-09-939-853A-97
/ Sequence 97, Application US/09939853A
/ Publication No. US20040039163A1
/ GENERAL INFORMATION:
/ APPLICANT: Burgess et al.
/ TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-099
/ CURRENT APPLICATION NUMBER: US/09/939,853A
/ CURRENT FILING DATE: 2001-08-27
/ PRIOR APPLICATION NUMBER: 60/228,191
/ PRIOR FILING DATE: 2000-08-25
/ PRIOR APPLICATION NUMBER: 60/267,300
/ PRIOR FILING DATE: 2001-02-08
/ PRIOR APPLICATION NUMBER: 60/269,961
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/277,337
/ PRIOR FILING DATE: 2001-03-20
/ NUMBER OF SEQ ID NOS: 159
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 97
/ LENGTH: 831
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-939-853A-97

Query Match      67.7%; Score 42; DB 3; Length 831;
Best Local Similarity 61.5%; Pred. No. 4.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      301 VTEENKELANELR 313

RESULT 31
US-10-296-733-1
/ Sequence 1, Application US/10296733
/ Publication No. US20040110131A1
/ GENERAL INFORMATION:
/ APPLICANT: Beth Israel Deaconess Medical Center
/ APPLICANT: Lawler, John W.
/ TITLE OF INVENTION: Thrombospondin-1 Type 1 Repeat
```

```
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 1440.2009002
/ CURRENT APPLICATION NUMBER: US/10/296,733
/ CURRENT FILING DATE: 2002-11-25
/ PRIOR APPLICATION NUMBER: 60/207,994
/ PRIOR FILING DATE: 2000-05-26
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1150
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-296-733-1

Query Match      67.7%; Score 42; DB 4; Length 1150;
Best Local Similarity 61.5%; Pred. No. 6.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      283 VTEENKELANELR 295

RESULT 32
US-09-919-603-1
/ Sequence 1, Application US/09919603
/ Patent No. US20020137679A1
/ GENERAL INFORMATION:
/ APPLICANT: Lawler, John W.
/ TITLE OF INVENTION: COMP/TSP-1, COMP/TSP-2 and Other TSP
/ FILE REFERENCE: 1440.1033-007
/ CURRENT APPLICATION NUMBER: US/09/919,603
/ CURRENT FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: PCT/US00/02482
/ PRIOR FILING DATE: 2000-02-01
/ PRIOR APPLICATION NUMBER: 60/118,053
/ PRIOR FILING DATE: 1999-02-01
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1152
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-919-603-1

Query Match      67.7%; Score 42; DB 3; Length 1152;
Best Local Similarity 61.5%; Pred. No. 6.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ILEENKELENALK 13
Db      283 VTEENKELANELR 295

RESULT 33
US-10-317-821B-7
/ Sequence 7, Application US/10317821B
/ Publication No. US20040241652A1
/ GENERAL INFORMATION:
/ APPLICANT: Aventis Pasteur Limited
/ TITLE OF INVENTION: Enhancement of the Immune Response Using CD-36 Binding Domain
/ FILE REFERENCE: API-01-17-PCT
/ CURRENT APPLICATION NUMBER: US/10/317,821B
/ CURRENT FILING DATE: 2002-12-12
/ PRIOR APPLICATION NUMBER: 60/341,771
/ PRIOR FILING DATE: 2001-12-12
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 7
/ LENGTH: 1169
/ TYPE: PRT
/ ORGANISM: Homo sapiens
```

US-10-317-821B-7

Query Match 67.7%; Score 42; DB 5; Length 1169;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 300 VTEENKELANEALR 312

RESULT 34

US-10-020-141-12
; Sequence 12, Application US/10020141
; Publication No. US20030092013A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Ableson, Allen
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-002
; CURRENT APPLICATION NUMBER: US/10/020,141
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-020-141-12

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 35

US-10-017-721-2
; Sequence 2, Application US/10017721
; Publication No. US20030096248A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Daley, George
; APPLICANT: Bolk, Stacey
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-003
; CURRENT APPLICATION NUMBER: US/10/017,721
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/317,033
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/330,248
; PRIOR FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-017-721-2

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 38

Db 301 VTEENKELANEALR 313

RESULT 36

US-10-021-660-114
; Sequence 114, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 114
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-021-660-114

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 37

US-10-008-093-2
; Sequence 2, Application US/10008093
; Publication No. US20030166017A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette Just
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE IDENTIFICATION,
; TITLE OF INVENTION: ASSESSMENT, PREVENTION AND THERAPY OF
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE
; FILE REFERENCE: MRI-026
; CURRENT APPLICATION NUMBER: US/10/008,093
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/248,185
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/257,417
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-008-093-2

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

```
US-10-295-027-1170
; Sequence 1170, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afaz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1170
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1170

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANEELR 313

RESULT 39
US-10-211-462-38
; Sequence 38, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
```

```
US-10-295-027-1170
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-38

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANEELR 313

RESULT 40
US-10-231-956A-482
; Sequence 482, Application US/10231956A
; Publication No. US2004005323A1
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Xu, Weiduan
; APPLICANT: Bogenberger, Jakob
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Modulators of Angiogenesis
; FILE REFERENCE: 021044-004100US
; CURRENT APPLICATION NUMBER: US/10/231,956A
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 482
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-956A-482

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANEELR 313

RESULT 41
US-10-419-462-38
; Sequence 38, Application US/10419462
; Publication No. US20040053392A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J.
; APPLICANT: Williams, Kevin J.
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof in Clinical Assays for
; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents
; FILE REFERENCE: W1107-20005
; CURRENT APPLICATION NUMBER: US/10/419,462
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
```

US-10-419-462-38

Query Match 67.7%; Score 42; DB 4; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 42

US-10-741-600-1018
; Sequence 1018, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1018
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1018

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 43

US-10-741-600-1019
; Sequence 1019, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1019
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1019

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 44

US-10-741-600-1021
; Sequence 1021, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1021
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1021

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 45

US-10-782-968-38
; Sequence 38, Application US/10782968
; Publication No. US20050065324A1
; GENERAL INFORMATION:
; APPLICANT: Kevin J. Williams
; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof in Clinical Assays for
; FILE REFERENCE: W107-20005
; CURRENT APPLICATION NUMBER: US/10/782,968
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/419,462
; PRIOR FILING DATE: 2003-04-21
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Human
US-10-782-968-38

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
Db 301 VTEENKELANELR 313

RESULT 46

US-10-849-989-44
; Sequence 44, Application US/10849989
; Publication No. US20050142569A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Guild, Braydon C.
; APPLICANT: Liao, Hua
; APPLICANT: Jones, Michael D.
; APPLICANT: Wu, Jiang
; APPLICANT: Zolg, Johannes W.
; TITLE OF INVENTION: Compositions, Kits, and Methods for
; TITLE OF INVENTION: Identification, Assessment, Prevention, and Therapy of
; TITLE OF INVENTION: Rheumatoid Arthritis
; FILE REFERENCE: MPI03-037PIRM
; CURRENT APPLICATION NUMBER: US/10/849,989
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: 60/472,330
; PRIOR FILING DATE: 2003-05-21
; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: human
US-10-849-989-44

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313

RESULT 47
US-10-631-467-548
; Sequence 548, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive p
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 548
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-548

Query Match 67.7%; Score 42; DB 5; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313

RESULT 48
US-11-037-713-51
; Sequence 51, Application US/11037713
; Publication No. US20050221398A1
; GENERAL INFORMATION:
; APPLICANT: JACQUEMIER, JOCELYNE
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: DEBONO, STEPHANE
; APPLICANT: TAGETT, REBECCA
; TITLE OF INVENTION: PROTEIN EXPRESSION PROFILING AND BREAST CANCER
; TITLE OF INVENTION: PROGNOSIS
; FILE REFERENCE: 1016-R-04(B)
; CURRENT APPLICATION NUMBER: US/11/037,713
; CURRENT FILING DATE: 2005-01-18
; PRIOR APPLICATION NUMBER: 60/537,412
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 51
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-037-713-51

Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 6.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313

RESULT 49
US-10-335-977-7129
; Sequence 7129, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 7129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 176 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...176
; SEQUENCE DESCRIPTION: SEQ ID NO: 7129:

Query Match 66.1%; Score 41; DB 4; Length 176;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 BENKELENA 11
Db 28 BENKELKNA 36

RESULT 50
US-11-097-143-2913
; Sequence 2913, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:

Mon Mar 13 10:54:38 2006

```
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2913
; LENGTH: 790
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2913
```

```
Query Match 66.1%; Score 41; DB 6; Length 790;
Best Local Similarity 61.5%; Pred. No. 6.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ILLENKELENALK 13
Db 32 LLERNKELETFIK 44
```

```
RESULT 51
US-09-815-242-11419
; Sequence 11419, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11577
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11577
```

```
Query Match 66.1%; Score 41; DB 3; Length 847;
Best Local Similarity 88.9%; Pred. No. 6.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 EENKELENA 11
Db 699 EENKELENA 707
```

```
RESULT 53
US-09-895-913A-164
; Sequence 164, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
```

```
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11419
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11419
```

```
Query Match 66.1%; Score 41; DB 3; Length 847;
Best Local Similarity 88.9%; Pred. No. 6.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 EENKELENA 11
Db 699 EENKELENA 707
```

```
RESULT 52
US-09-815-242-11577
; Sequence 11577, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11577
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11577
```

```
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-895-913A-164

Query Match 66.1%; Score 41; DB 3; Length 847;
Best Local Similarity 88.9%; Pred. No. 6.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EENKELENA 11
Db 699 EENKELKNA 707

RESULT 54
US-10-282-122A-58924
; Sequence 58924, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58924
```

```
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-282-122A-58924

Query Match 66.1%; Score 41; DB 4; Length 847;
Best Local Similarity 88.9%; Pred. No. 6.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EENKELENA 11
Db 699 EENKELKNA 707

RESULT 55
US-10-335-977-7131
; Sequence 7131, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7131:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...847
; SEQUENCE DESCRIPTION: SEQ ID NO: 7131:
US-10-335-977-7131

Query Match 66.1%; Score 41; DB 4; Length 847;
Best Local Similarity 88.9%; Pred. No. 6.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EENKELENA 11
Db 699 EENKELKNA 707
```

```
RESULT 56
US-10-036-869-1
; Sequence 1, Application US/10036869
; Publication No. US20020151516A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/036,869
; FILING DATE: 29-NO. US20020151516A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-036-869-1
Query Match 64.5%; Score 40; DB 4; Length 218;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
||||| |
DB 3 EENKELANELR 13

RESULT 57
US-10-205-194-99
; Sequence 99, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
; LENGTH: 245
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: BAP31
US-10-205-194-99
Query Match 64.5%; Score 40; DB 4; Length 245;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
||||| |
DB 177 LEENKSLKNDLR 188

RESULT 58
US-10-767-701-42480
; Sequence 42480, Application US/10767701
; Publication No. US20040172884A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42480
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CS2891_1.pep
US-10-767-701-42480
Query Match 64.5%; Score 40; DB 4; Length 275;
Best Local Similarity 72.7%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENALK 13
||||| |
DB 197 BEGQLENRLK 207

RESULT 59
US-10-425-115-226060
; Sequence 226060, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 226060
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_137756C.1.pep
US-10-425-115-226060
Query Match 64.5%; Score 40; DB 4; Length 386;
Best Local Similarity 72.7%; Pred. No. 4.1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```


Mon Mar 13 10:54:38 2006

QY 3 BENKELENALK 13
|||:|||||
Db 309 BEGQLENRLK 319

RESULT 60

US-09-815-242-13407
; Sequence 13407, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13407
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13407

Query Match 64.5%; Score 40; DB 3; Length 391;
Best Local Similarity 71.4%; Pred. NO. 4.1e+02;
Matches 10; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 2 LEENKE--LENALK 13
|||||:|||||
Db 269 LEENREKSLKALK 282

Search completed: March 11, 2006, 12:32:33
Job time : 109.947 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2006, 12:27:37 ; Search time 11.9737 Seconds
(without alignments)
30.221 Million cell updates/sec

Title: US-10-774-242A-6
Perfect score: 62
Sequence: 1 ILEENKELENALK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 500 summaries

Database: Published Applications AA.New.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	760	7	US-11-186-284-55
2	42	67.7	1170	6	US-10-831-997-2
3	42	67.7	1170	6	US-10-995-561-594
4	42	67.7	1170	6	US-10-995-561-595
5	42	67.7	1170	6	US-10-995-561-596
6	42	67.7	1170	6	US-11-046-456-28
7	42	67.7	1170	7	US-11-046-644-28
8	41	66.1	291	7	US-11-087-099-12226
9	41	66.1	291	7	US-11-096-568A-29720
10	41	66.1	298	7	US-11-087-099-5401
11	38	61.3	51	7	US-11-219-282-31
12	38	61.3	430	6	US-10-525-710-30
13	37	59.7	192	7	US-11-096-568A-8540
14	37	59.7	208	7	US-11-096-568A-8539
15	37	59.7	222	7	US-11-096-568A-8538
16	37	59.7	246	7	US-11-096-568A-8537
17	37	59.7	284	7	US-11-096-568A-4493
18	37	59.7	308	7	US-11-096-568A-4494
19	37	59.7	310	7	US-11-087-099-2317
20	37	59.7	421	6	US-10-525-710-42
21	36	58.1	289	7	US-11-087-099-8567
22	36	58.1	330	7	US-11-087-099-10594
23	36	58.1	334	7	US-11-087-099-8967
24	36	58.1	582	6	US-10-763-712A-23
25	36	58.1	582	6	US-10-763-712A-101

245	31	50.0	340	7	US-11-096-568A-27013	Sequence 27013, A	318	30.5	49.2	454	7	US-11-098-686-10919	Sequence 10919, A
246	31	50.0	348	6	US-10-793-626-1702	Sequence 1702, App	319	30.5	49.2	622	6	US-10-053-877-239	Sequence 239, App
247	31	50.0	349	7	US-11-018-868-118	Sequence 118, App	320	30.5	49.2	635	6	US-10-055-877-67	Sequence 67, Appl
248	31	50.0	351	6	US-10-793-626-338	Sequence 338, App	321	30.5	49.2	701	6	US-10-055-877-65	Sequence 65, Appl
249	31	50.0	351	7	US-11-096-568A-18664	Sequence 18664, A	322	30.5	49.2	885	6	US-10-055-877-240	Sequence 240, App
250	31	50.0	353	7	US-11-096-568A-27147	Sequence 27147, A	323	30.5	49.2	885	6	US-10-055-877-241	Sequence 241, App
251	31	50.0	357	7	US-11-096-568A-15042	Sequence 15042, A	324	30	48.4	105	7	US-11-000-463-818	Sequence 818, App
252	31	50.0	370	7	US-11-096-568A-11951	Sequence 11951, A	325	30	48.4	111	6	US-10-667-295-145	Sequence 145, App
253	31	50.0	378	7	US-11-096-568A-11951	Sequence 11951, A	326	30	48.4	111	6	US-11-096-568A-3696	Sequence 3696, App
254	31	50.0	387	7	US-11-096-568A-11950	Sequence 11950, A	327	30	48.4	114	6	US-10-793-626-2252	Sequence 2252, App
255	31	50.0	388	7	US-11-096-568A-3895	Sequence 3895, App	328	30	48.4	120	6	US-10-667-295-144	Sequence 144, App
256	31	50.0	404	7	US-11-096-568A-11949	Sequence 11949, A	329	30	48.4	120	6	US-11-096-568A-3695	Sequence 3695, App
257	31	50.0	404	7	US-11-096-568A-12273	Sequence 12273, A	330	30	48.4	127	6	US-10-793-626-2060	Sequence 2060, App
258	31	50.0	406	6	US-10-821-234-1521	Sequence 1521, App	331	30	48.4	133	6	US-11-176-830-454	Sequence 454, App
259	31	50.0	411	7	US-11-096-568A-13272	Sequence 13272, A	332	30	48.4	135	6	US-10-821-234-1018	Sequence 1018, App
260	31	50.0	413	6	US-10-517-939-110	Sequence 110, App	333	30	48.4	146	7	US-11-087-099-648	Sequence 648, App
261	31	50.0	416	7	US-11-096-568A-27012	Sequence 27012, A	334	30	48.4	153	6	US-10-467-657-6118	Sequence 6118, App
262	31	50.0	421	7	US-11-096-568A-31682	Sequence 31682, A	335	30	48.4	158	7	US-11-096-568A-13852	Sequence 13852, A
263	31	50.0	422	7	US-11-096-568A-11396	Sequence 11396, App	336	30	48.4	160	7	US-11-096-568A-13851	Sequence 13851, A
264	31	50.0	426	7	US-11-098-686-11396	Sequence 11396, A	337	30	48.4	166	6	US-10-821-234-1293	Sequence 1293, App
265	31	50.0	440	7	US-11-096-568A-31383	Sequence 31383, A	338	30	48.4	166	6	US-10-878-556A-20	Sequence 20, Appl
266	31	50.0	440	7	US-11-096-568A-31681	Sequence 31681, A	339	30	48.4	166	7	US-11-177-506-33	Sequence 33, Appl
267	31	50.0	443	6	US-10-793-626-1200	Sequence 1200, App	340	30	48.4	173	7	US-11-096-568A-9755	Sequence 9755, App
268	31	50.0	447	7	US-11-096-568A-12271	Sequence 12271, A	341	30	48.4	180	7	US-11-098-686-10202	Sequence 10202, A
269	31	50.0	449	7	US-11-177-506-32	Sequence 32, Appl	342	30	48.4	194	7	US-11-108-172-126	Sequence 1126, App
270	31	50.0	455	6	US-10-772-512-3400	Sequence 3400, App	343	30	48.4	195	7	US-11-096-568A-9754	Sequence 13850, A
271	31	50.0	463	6	US-10-821-234-1094	Sequence 1094, App	344	30	48.4	205	7	US-11-096-568A-13850	Sequence 13850, A
272	31	50.0	476	7	US-11-124-368A-261	Sequence 261, App	345	30	48.4	208	6	US-10-793-626-782	Sequence 782, App
273	31	50.0	486	7	US-11-096-568A-27011	Sequence 27011, A	346	30	48.4	208	7	US-11-129-104-97	Sequence 97, Appl
274	31	50.0	504	7	US-11-072-512-3467	Sequence 3467, App	347	30	48.4	208	7	US-11-087-099-5674	Sequence 5674, App
275	31	50.0	522	7	US-11-096-568A-18662	Sequence 18662, A	348	30	48.4	222	6	US-10-667-295-143	Sequence 143, App
276	31	50.0	529	7	US-11-096-568A-31382	Sequence 31382, A	349	30	48.4	222	7	US-11-124-367A-269	Sequence 269, App
277	31	50.0	544	6	US-10-467-657-1020	Sequence 1020, App	350	30	48.4	222	7	US-11-072-512-3044	Sequence 3044, App
278	31	50.0	545	7	US-11-201-916-19	Sequence 19, Appl	351	30	48.4	222	7	US-11-096-568A-3694	Sequence 3694, App
279	31	50.0	548	7	US-11-096-568A-26924	Sequence 26924, A	352	30	48.4	227	7	US-11-087-099-3662	Sequence 3662, App
280	31	50.0	548	7	US-11-096-568A-27017	Sequence 27017, A	353	30	48.4	227	7	US-11-087-099-4634	Sequence 4634, App
281	31	50.0	576	6	US-10-530-340-12	Sequence 12, Appl	354	30	48.4	241	6	US-10-485-517-189	Sequence 189, App
282	31	50.0	604	7	US-11-096-568A-27016	Sequence 27016, A	355	30	48.4	241	7	US-11-096-568A-6296	Sequence 6296, App
283	31	50.0	610	7	US-11-096-568A-27015	Sequence 27015, A	356	30	48.4	282	7	US-11-087-099-4082	Sequence 4082, App
284	31	50.0	610	7	US-11-096-568A-31381	Sequence 31381, A	357	30	48.4	282	7	US-11-087-099-7056	Sequence 7056, App
285	31	50.0	619	7	US-11-124-368A-260	Sequence 260, App	358	30	48.4	282	7	US-11-087-099-10507	Sequence 10507, A
286	31	50.0	636	6	US-10-485-517-170	Sequence 170, App	359	30	48.4	283	7	US-11-087-099-3645	Sequence 3645, App
287	31	50.0	655	7	US-11-124-368A-262	Sequence 262, App	360	30	48.4	283	7	US-11-087-099-12351	Sequence 12351, A
288	31	50.0	663	7	US-11-096-568A-26923	Sequence 26923, A	361	30	48.4	293	6	US-10-000-997-30	Sequence 30, Appl
289	31	50.0	671	7	US-11-124-368A-259	Sequence 259, App	362	30	48.4	298	7	US-11-087-099-10462	Sequence 10462, A
290	31	50.0	683	6	US-10-821-234-1314	Sequence 1314, App	363	30	48.4	300	7	US-11-087-099-8453	Sequence 8453, App
291	31	50.0	683	7	US-11-081-566-2	Sequence 2, Appl	364	30	48.4	303	6	US-10-467-9628-15	Sequence 15, Appl
292	31	50.0	683	7	US-11-108-172-122	Sequence 122, App	365	30	48.4	303	6	US-11-087-099-8877	Sequence 8877, App
293	31	50.0	683	7	US-11-186-284-201	Sequence 201, App	366	30	48.4	310	7	US-11-010-239-97	Sequence 97, Appl
294	31	50.0	699	7	US-11-096-568A-26922	Sequence 26922, A	367	30	48.4	313	7	US-11-096-568A-6295	Sequence 6295, App
295	31	50.0	737	7	US-11-152-366-28	Sequence 28, Appl	368	30	48.4	321	7	US-11-264-728-40	Sequence 40, Appl
296	31	50.0	757	7	US-11-097-463-2	Sequence 2, Appl	369	30	48.4	321	7	US-11-087-099-11191	Sequence 11191, A
297	31	50.0	840	7	US-11-207-078-190	Sequence 190, App	370	30	48.4	329	7	US-11-087-099-2753	Sequence 2753, App
298	31	50.0	844	7	US-11-097-463-4	Sequence 4, Appl	371	30	48.4	334	7	US-11-096-568A-6294	Sequence 6294, App
299	31	50.0	868	7	US-11-096-568A-31321	Sequence 31321, A	372	30	48.4	340	7	US-11-087-099-4507	Sequence 4507, App
300	31	50.0	872	7	US-11-207-078-221	Sequence 221, App	373	30	48.4	343	7	US-11-087-099-11420	Sequence 11420, A
301	31	50.0	876	7	US-11-096-568A-31320	Sequence 31320, A	374	30	48.4	345	7	US-11-087-099-5483	Sequence 5483, App
302	31	50.0	884	7	US-11-098-686-11426	Sequence 11426, A	375	30	48.4	349	7	US-11-087-099-11218	Sequence 11218, A
303	31	50.0	937	7	US-11-096-568A-31319	Sequence 31319, A	376	30	48.4	358	7	US-11-087-099-8631	Sequence 8631, App
304	31	50.0	1014	6	US-10-501-035-307	Sequence 307, App	377	30	48.4	362	6	US-11-087-099-9573	Sequence 9573, App
305	31	50.0	1280	7	US-11-045-578-4	Sequence 4, Appl	378	30	48.4	363	6	US-10-793-626-302	Sequence 302, App
306	31	50.0	1404	6	US-10-878-556A-169	Sequence 169, App	379	30	48.4	364	7	US-11-096-568A-5018	Sequence 5018, App
307	31	50.0	1728	7	US-11-019-711-72	Sequence 72, Appl	380	30	48.4	366	7	US-11-087-099-2427	Sequence 2427, App
308	31	50.0	1927	7	US-11-087-099-5472	Sequence 5472, App	381	30	48.4	366	7	US-11-087-099-11191	Sequence 11191, A
309	31	50.0	1971	7	US-11-179-624-1	Sequence 1, Appl	382	30	48.4	369	6	US-10-763-712A-43	Sequence 43, Appl
310	31	50.0	1980	7	US-11-179-624-3	Sequence 3, Appl	383	30	48.4	383	7	US-11-072-512-3412	Sequence 3412, App
311	31	50.0	2455	7	US-11-186-999-14	Sequence 14, Appl	384	30	48.4	385	6	US-11-051-720-1348	Sequence 1348, App
312	31	50.0	2455	7	US-11-186-999-16	Sequence 16, Appl	385	30	48.4	389	6	US-10-763-712A-31	Sequence 31, Appl
313	31	50.0	2456	7	US-11-186-999-8	Sequence 8, Appl	386	30	48.4	403	7	US-11-072-512-2845	Sequence 2845, App
314	31	50.0	2456	7	US-11-186-999-10	Sequence 10, Appl	387	30	48.4	411	7	US-11-072-512-3452	Sequence 3452, App
315	30.5	49.2	135	7	US-11-087-099-3769	Sequence 3769, App	388	30	48.4	413	6	US-10-821-234-989	Sequence 989, App
316	30.5	49.2	358	7	US-11-087-099-2906	Sequence 2906, App	389	30	48.4	416	7	US-11-096-568A-3578	Sequence 3578, App
317	30.5	49.2	364	7	US-11-087-099-9259	Sequence 9259, App	390	30	48.4	424	7	US-11-096-568A-3577	Sequence 3577, App

391	30	48.4	424	7	US-11-096-568A-5017	Sequence 5017, Ap	464	30	48.4	770	6	US-10-821-234-1269	Sequence 1269, Ap
392	30	48.4	425	6	US-10-467-657-8052	Sequence 8052, Ap	465	30	48.4	777	7	US-11-124-368A-238	Sequence 238, App
393	30	48.4	426	7	US-11-096-568A-30753	Sequence 30753, A	466	30	48.4	778	7	US-11-124-368A-240	Sequence 240, App
394	30	48.4	436	7	US-11-087-099-11618	Sequence 11618, A	467	30	48.4	791	7	US-11-124-368A-245	Sequence 245, App
395	30	48.4	436	7	US-11-096-568A-7059	Sequence 7059, Ap	468	30	48.4	803	7	US-11-124-368A-241	Sequence 241, App
396	30	48.4	437	6	US-10-453-372-438	Sequence 438, App	469	30	48.4	803	7	US-11-124-368A-242	Sequence 242, App
397	30	48.4	438	7	US-11-096-568A-5016	Sequence 5016, Ap	470	30	48.4	814	7	US-11-096-568A-28352	Sequence 28352, A
398	30	48.4	449	7	US-11-096-568A-7058	Sequence 7058, Ap	471	30	48.4	816	7	US-11-096-568A-28351	Sequence 28351, A
399	30	48.4	458	7	US-11-096-568A-7057	Sequence 7057, Ap	472	30	48.4	831	7	US-11-096-568A-1542	Sequence 1542, Ap
400	30	48.4	459	7	US-11-087-099-8021	Sequence 8021, Ap	473	30	48.4	836	7	US-11-124-368A-246	Sequence 246, App
401	30	48.4	459	7	US-11-087-099-10708	Sequence 10708, A	474	30	48.4	858	7	US-11-054-281-135	Sequence 135, App
402	30	48.4	463	7	US-11-087-099-4304	Sequence 4304, Ap	475	30	48.4	867	7	US-11-054-281-134	Sequence 134, App
403	30	48.4	468	7	US-11-096-568A-32982	Sequence 32982, A	476	30	48.4	911	6	US-10-330-773-712	Sequence 712, App
404	30	48.4	470	7	US-11-072-512-2022	Sequence 2022, Ap	477	30	48.4	957	7	US-11-051-720-1438	Sequence 1438, Ap
405	30	48.4	475	6	US-10-330-773-227	Sequence 227, App	478	30	48.4	1011	7	US-11-089-642-111	Sequence 111, App
406	30	48.4	489	6	US-10-793-626-1302	Sequence 1302, Ap	479	30	48.4	1031	6	US-10-857-780-22	Sequence 22, Appl
407	30	48.4	490	7	US-11-183-372-23	Sequence 23, Appl	480	30	48.4	1041	6	US-10-828-831-9	Sequence 9, Appl
408	30	48.4	492	7	US-11-072-512-3704	Sequence 3704, Ap	481	30	48.4	1059	7	US-11-231-599-10	Sequence 10, Appl
409	30	48.4	501	7	US-11-087-099-7328	Sequence 7328, Ap	482	30	48.4	1076	7	US-11-096-568A-29101	Sequence 29101, A
410	30	48.4	505	6	US-10-763-712A-26	Sequence 26, Appl	483	30	48.4	1082	7	US-11-096-568A-29100	Sequence 29100, A
411	30	48.4	505	6	US-10-763-712A-88	Sequence 88, Appl	484	30	48.4	1086	7	US-11-082-554A-380	Sequence 380, App
412	30	48.4	505	6	US-10-763-712A-89	Sequence 89, Appl	485	30	48.4	1114	6	US-10-857-780-27	Sequence 27, Appl
413	30	48.4	513	6	US-10-453-372-442	Sequence 442, App	486	30	48.4	1151	6	US-10-793-626-2448	Sequence 2448, Ap
414	30	48.4	514	7	US-11-051-720-1346	Sequence 1346, Ap	487	30	48.4	1155	6	US-10-755-092-9	Sequence 9, Appl
415	30	48.4	520	7	US-11-096-568A-30752	Sequence 30752, A	488	30	48.4	1181	6	US-10-755-092-11	Sequence 11, Appl
416	30	48.4	521	7	US-11-087-099-6169	Sequence 6169, Ap	489	30	48.4	1181	6	US-10-755-092-13	Sequence 13, Appl
417	30	48.4	535	6	US-10-453-372-436	Sequence 436, App	490	30	48.4	1181	6	US-10-755-092-15	Sequence 15, Appl
418	30	48.4	535	6	US-10-493-909-84	Sequence 84, Appl	491	30	48.4	1181	6	US-10-755-092-17	Sequence 17, Appl
419	30	48.4	542	7	US-11-096-568A-30751	Sequence 30751, A	492	30	48.4	1237	7	US-11-052-554A-95	Sequence 95, Appl
420	30	48.4	543	6	US-10-330-773-224	Sequence 224, App	493	30	48.4	1327	7	US-11-096-568A-29099	Sequence 29099, A
421	30	48.4	551	7	US-11-000-463-346	Sequence 346, App	494	30	48.4	1347	7	US-11-087-059-5370	Sequence 5370, A
422	30	48.4	557	6	US-11-264-728-36	Sequence 36, Appl	495	30	48.4	1388	6	US-10-821-234-1143	Sequence 1143, Ap
423	30	48.4	567	6	US-10-420-192-8	Sequence 8, Appl	496	30	48.4	1388	6	US-10-501-035-350	Sequence 350, App
424	30	48.4	567	6	US-11-072-175-197	Sequence 197, App	497	30	48.4	1452	6	US-10-985-561-778	Sequence 778, App
425	30	48.4	573	6	US-10-467-657-684	Sequence 684, App	498	30	48.4	1452	6	US-11-231-599-8	Sequence 8, Appl
426	30	48.4	574	7	US-11-096-568A-32981	Sequence 32981, A	499	30	48.4	1628	7	US-10-055-877-147	Sequence 147, App
427	30	48.4	586	7	US-11-072-512-2947	Sequence 2947, Ap	500	30	48.4	1732	6		
428	30	48.4	586	7	US-11-072-512-3485	Sequence 3485, Ap							
429	30	48.4	602	7	US-11-087-099-11542	Sequence 11542, A							
430	30	48.4	606	7	US-11-072-512-3888	Sequence 3888, Ap							
431	30	48.4	610	7	US-11-194-246-313	Sequence 313, App							
432	30	48.4	625	7	US-11-055-822-522	Sequence 522, App							
433	30	48.4	629	7	US-11-096-568A-32980	Sequence 32980, A							
434	30	48.4	641	7	US-11-128-660-1	Sequence 28353, A							
435	30	48.4	651	7	US-10-821-234-1477	Sequence 1477, Ap							
436	30	48.4	667	6	US-11-124-368A-250	Sequence 250, App							
437	30	48.4	674	7	US-11-051-720-1342	Sequence 1342, Ap							
438	30	48.4	676	6	US-10-453-372-440	Sequence 440, App							
439	30	48.4	683	7	US-11-124-368A-243	Sequence 243, App							
440	30	48.4	686	7	US-11-124-368A-239	Sequence 239, App							
441	30	48.4	691	6	US-10-467-657-7170	Sequence 7170, Ap							
442	30	48.4	708	7	US-11-072-175-158	Sequence 158, App							
443	30	48.4	714	7	US-11-124-367A-268	Sequence 268, App							
444	30	48.4	725	7	US-11-051-720-1341	Sequence 1341, Ap							
445	30	48.4	726	7	US-11-124-368A-247	Sequence 247, App							
446	30	48.4	730	6	US-10-453-372-444	Sequence 444, App							
447	30	48.4	732	6	US-11-124-368A-248	Sequence 248, App							
448	30	48.4	737	6	US-10-453-372-434	Sequence 434, App							
449	30	48.4	737	6	US-10-453-372-446	Sequence 446, App							
450	30	48.4	737	6	US-10-453-372-448	Sequence 448, App							
451	30	48.4	737	6	US-10-453-372-450	Sequence 450, App							
452	30	48.4	737	6	US-10-453-372-452	Sequence 452, App							
453	30	48.4	737	6	US-10-453-372-454	Sequence 454, App							
454	30	48.4	737	6	US-10-453-372-456	Sequence 456, App							
455	30	48.4	737	6	US-11-051-720-1345	Sequence 1345, Ap							
456	30	48.4	738	6	US-10-595-561-692	Sequence 692, App							
457	30	48.4	738	6	US-10-828-831-5	Sequence 693, App							
458	30	48.4	746	6	US-10-828-831-7	Sequence 5, Appl							
459	30	48.4	746	6	US-11-124-368A-244	Sequence 244, App							
460	30	48.4	750	7	US-11-124-368A-249	Sequence 249, App							
461	30	48.4	753	7	US-11-124-368A-237	Sequence 237, App							
462	30	48.4											
463	30	48.4											

RESULT 1
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US2005026493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P2RNM
; CURRENT FILING DATE: 2005-07-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-12-10
; PRIOR FILING DATE: 2002-03-05
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760

ALIGNMENTS

TYPE: PRT
ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match 100.0%; Score 62; DB 7; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.018; 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 487 ILEENKELENALK 499

RESULT 2

US-10-831-997-2
; Sequence 2, Application US/10831997
; Publication No. US2005024834A1
; GENERAL INFORMATION:
; APPLICANT: Lander, Eric S.
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Bolk, Stacey
; APPLICANT: Daley, George Q.
; APPLICANT: McCarthy, Jeanette J.
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
; FILE REFERENCE: 2825.1027-001
; CURRENT APPLICATION NUMBER: US/10/831,997
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: US/09/657,472
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/153,357
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/220,947
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US 60/225,724
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2551
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-831-997-2

Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52; 3; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 3

US-10-995-561-594
; Sequence 594, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-594

Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52; 3; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 4

US-10-995-561-595
; Sequence 595, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 595
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-595

Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52; 3; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 5

US-10-995-561-596
; Sequence 596, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 596
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-596

Query Match 67.7%; Score 42; DB 6; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52; 3; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENALK 13
: ||||| :
DB 301 VTEENKELANEALR 313

RESULT 6

US-11-046-456-28
; Sequence 28, Application US/11046456
; Publication No. US2005025604A1
; GENERAL INFORMATION:
; APPLICANT: Boyle, Bryan

```
; APPLICANT: Funk, Walter
; APPLICANT: Kakitani, Makoto
; APPLICANT: Oshima, Takeshi
; APPLICANT: Park, Eun Ju
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yagi, Mikio
; APPLICANT: Tomizuka, Kazuma
; TITLE OF INVENTION: Gastrointestinal Proliferative Factor and uses thereof
; FILE REFERENCE: 18824-003001
; CURRENT APPLICATION NUMBER: US 11/046,456
; CURRENT FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US 60/539,605
; PRIOR FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 60/619,241
; PRIOR FILING DATE: 2004-10-15
; PRIOR APPLICATION NUMBER: US 10/488,423
; PRIOR FILING DATE: 2004-03-03
; PRIOR APPLICATION NUMBER: PCT/US02/27746
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-046-644-28
```

```
Query Match 67.7%; Score 42; DB 7; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
```

```
RESULT 7
US-11-046-644-28
; Sequence 28, Application US/11046644
; Publication No. US20050256036A1
; GENERAL INFORMATION:
; APPLICANT: Boyle, Bryan
; APPLICANT: Funk, Walter
; APPLICANT: Kakitani, Makoto
; APPLICANT: Oshima, Takeshi
; APPLICANT: Park, Emily
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tomizuka, Kazuma
; APPLICANT: Yagi, Mikio
; TITLE OF INVENTION: Gastrointestinal Proliferative Factor and uses thereof
; FILE REFERENCE: 11926-194001
; CURRENT APPLICATION NUMBER: US/11/046,644
; CURRENT FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US 60/539,605
; PRIOR FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 60/619,241
; PRIOR FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-046-644-28
```

```
Query Match 67.7%; Score 42; DB 7; Length 1170;
Best Local Similarity 61.5%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 ILEENKELENALK 13
Db 301 VTEENKELANELR 313
```

```
RESULT 8
US-11-087-099-12226
; Sequence 12226, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 12226
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-12226
```

```
Query Match 66.1%; Score 41; DB 7; Length 291;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 3 EENKELENAL 12
Db 25 EENKKEFENAL 34
```

```
RESULT 9
US-11-096-568A-29720
; Sequence 29720, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29720
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(291)
; OTHER INFORMATION: Ceres Seq. ID no. 4926980
US-11-096-568A-29720
```

```
Query Match 66.1%; Score 41; DB 7; Length 291;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 3 EENKELENAL 12
Db 25 EENKKEFENAL 34
```

```
RESULT 10
US-11-087-099-5401
; Sequence 5401, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5401
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
```

US-11-087-099-5401

Query Match 66.1%; Score 41; DB 7; Length 298;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 ILEENKELENA 12
DB 32 ILEENKELENA 41

RESULT 11

US-11-219-282-31
; Sequence 31, Application US/11219282
; Publication No. US20060018892A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: DuBose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/11/219,282
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US/10/094,944
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-219-282-31

Query Match 61.3%; Score 38; DB 7; Length 51;
Best Local Similarity 63.6%; Pred. No. 5.9;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ILEENKELENA 12
DB 31 ILEENKELENA 41

RESULT 12

US-10-525-710-30
; Sequence 30, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolpprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; TITLE OF INVENTION: Sulphur (met)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-525-710-30

Query Match 61.3%; Score 38; DB 6; Length 430;
Best Local Similarity 69.2%; Pred. No. 72;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ILEENKELENA 13
DB 322 ILEENKELENA 334

RESULT 13

US-11-096-568A-8540
; Sequence 8540, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8540
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)-(192)
; OTHER INFORMATION: Ceres Seq. ID no. 15227122
US-11-096-568A-8540

Query Match 59.7%; Score 37; DB 7; Length 192;
Best Local Similarity 53.8%; Pred. No. 40;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENA 13
DB 165 ILEENKELENA 177

RESULT 14

US-11-096-568A-8539
; Sequence 8539, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8539
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)-(208)
; OTHER INFORMATION: Ceres Seq. ID no. 15227121
US-11-096-568A-8539

Query Match 59.7%; Score 37; DB 7; Length 208;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENA 13
DB 181 ILEENKELENA 193

RESULT 15

US-11-096-568A-8538
; Sequence 8538, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:


```
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8538
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(222)
; OTHER INFORMATION: Ceres Seq. ID no. 15227120
US-11-096-568A-8538

Query Match          59.7%; Score 37; DB 7; Length 222;
Best Local Similarity 53.8%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKLEENALK 13
Db 195 LLEENQNLRLNQLQ 207

RESULT 16
US-11-096-568A-8537
; Sequence 8537, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8537
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(246)
; OTHER INFORMATION: Ceres Seq. ID no. 15227119
US-11-096-568A-8537

Query Match          59.7%; Score 37; DB 7; Length 246;
Best Local Similarity 53.8%; Pred. No. 54;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKLEENALK 13
Db 219 LLEENQNLRLNQLQ 231

RESULT 17
US-11-096-568A-4494
; Sequence 4494, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4494
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Glycine max
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(284)
; OTHER INFORMATION: Ceres Seq. ID no. 15220412
US-11-096-568A-4494

Query Match          59.7%; Score 37; DB 7; Length 284;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKLEENAL 12
Db 14 EENKLEENAL 23

RESULT 18
US-11-096-568A-4493
; Sequence 4493, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4493
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(308)
; OTHER INFORMATION: Ceres Seq. ID no. 15220411
US-11-096-568A-4493

Query Match          59.7%; Score 37; DB 7; Length 308;
Best Local Similarity 80.0%; Pred. No. 70;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKLEENAL 12
Db 38 EENKLEENAL 47

RESULT 19
US-11-087-099-2317
; Sequence 2317, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2317
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Brassica rapa subsp. pekinensis
US-11-087-099-2317

Query Match          59.7%; Score 37; DB 7; Length 310;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKLEENALK 13
Db 74 LEKNFELENKLE 85

RESULT 20
```

Mon Mar 13 10:54:38 2006

```
US-10-525-710-42
; Sequence 42, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE OF INVENTION: Sulphur (metY)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 42
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-525-710-42

Query Match          59.7%; Score 37; DB 6; Length 421;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LEENKELENALK 13
      ||.|||
Db      318 LEDNKYENAKK 329

RESULT 21
US-11-087-099-8567
; Sequence 8567, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8567
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(289)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-8567

Query Match          58.1%; Score 36; DB 7; Length 289;
Best Local Similarity 70.0%; Pred. No. 95;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 EENKELENAL 12
      |||||
Db      31 EENKQFERAL 40

RESULT 22
US-11-087-099-10594
; Sequence 10594, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
US-10-525-710-42
; Sequence 42, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE OF INVENTION: Sulphur (metY)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 42
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-525-710-42

Query Match          59.7%; Score 37; DB 6; Length 421;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LEENKELENALK 13
      ||.|||
Db      318 LEDNKYENAKK 329

RESULT 21
US-11-087-099-8567
; Sequence 8567, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8567
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(289)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-8567

Query Match          58.1%; Score 36; DB 7; Length 289;
Best Local Similarity 70.0%; Pred. No. 95;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 EENKELENAL 12
      |||||
Db      31 EENKQFERAL 40

RESULT 22
US-11-087-099-10594
; Sequence 10594, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
US-10-525-710-42
; Sequence 42, Application US/10525710
; Publication No. US20050260721A1
; GENERAL INFORMATION:
; APPLICANT: Kroger, Burkhard
; APPLICANT: Zelder, Oskar
; APPLICANT: Kolprogge, Corinna
; APPLICANT: Schroder, Hartwig
; APPLICANT: Hafner, Stefan
; TITLE OF INVENTION: Method for Zymotic Production of Fine Chemicals Containing
; FILE OF INVENTION: Sulphur (metY)
; FILE REFERENCE: 13111-00006-US
; CURRENT APPLICATION NUMBER: US/10/525,710
; CURRENT FILING DATE: 2005-02-24
; PRIOR APPLICATION NUMBER: PCT/EP 2003/009453
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: DE 102 39 082.7
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 42
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Clostridium difficile
US-10-525-710-42

Query Match          59.1%; Score 36; DB 7; Length 330;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 EENKELENAL 12
      |||||
Db      31 EENKQFERAL 40

RESULT 23
US-11-087-099-8967
; Sequence 8967, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8967
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Zea mays
US-11-087-099-8967

Query Match          58.1%; Score 36; DB 7; Length 334;
Best Local Similarity 72.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 LEENKELENAL 12
      |||||
Db      66 LEENKQFERAL 76

RESULT 24
US-10-763-712A-23
; Sequence 23, Application US/10763712A
; Publication No. US20050266541A1
; GENERAL INFORMATION:
; APPLICANT: Solazyme, Inc.
; APPLICANT: Dillon, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Microbial Hydrogen
; TITLE OF INVENTION: Production
; FILE REFERENCE: H2042101-CIP
; CURRENT APPLICATION NUMBER: US/10/763,712A
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 10/287,750
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: US 10/411,910
; PRIOR FILING DATE: 2003-04-12
; PRIOR APPLICATION NUMBER: US 60/500,032
; PRIOR FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Trichomonas vaginalis
US-10-763-712A-23

Query Match          58.1%; Score 36; DB 6; Length 582;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 4 ENKELEN 10
Db 435 ENKELEN 441

RESULT 25

US-10-763-712A-101
; Sequence 101, Application US/10763712A
; Publication No. US20050266541A1
; GENERAL INFORMATION:
; APPLICANT: Solazyme, Inc.
; TITLE OF INVENTION: Methods and Compositions for Evolving Microbial Hydrogen
; FILE REFERENCE: H2042101-CIP
; CURRENT APPLICATION NUMBER: US 10/763,712A
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 10/287,750
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: US 10/411,910
; PRIOR FILING DATE: 2003-04-12
; PRIOR APPLICATION NUMBER: US 60/500,032
; PRIOR FILING DATE: 2003-09-03
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 101
; LENGTH: 582
; TYPE: PRT
; ORGANISM: C. acetobutylicum
US-10-763-712A-101

Query Match 58.1%; Score 36; DB 6; Length 582;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ENKELEN 10
Db 435 ENKELEN 441

RESULT 26

US-11-096-568A-19858
; Sequence 19858, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19858
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(729)
; OTHER INFORMATION: Ceres Seq. ID no. 12375068
US-11-096-568A-19858

Query Match 58.1%; Score 36; DB 7; Length 729;
Best Local Similarity 46.2%; Pred. No. 2.8e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 488 VMEETKEKNAVE 500

RESULT 27

US-11-096-568A-19857
; Sequence 19857, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19857
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(791)
; OTHER INFORMATION: Ceres Seq. ID no. 12375067
US-11-096-568A-19857

Query Match 58.1%; Score 36; DB 7; Length 791;
Best Local Similarity 46.2%; Pred. No. 3.1e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 550 VMEETKEKNAVE 562

RESULT 28

US-11-096-568A-19856
; Sequence 19856, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19856
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(848)
; OTHER INFORMATION: Ceres Seq. ID no. 12375066
US-11-096-568A-19856

Query Match 58.1%; Score 36; DB 7; Length 848;
Best Local Similarity 46.2%; Pred. No. 3.3e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 607 VMEETKEKNAVE 619

RESULT 29

US-11-054-281-116
; Sequence 116, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11

Mon Mar 13 10:54:38 2006

```
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 116
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-054-281-116

Query Match      58.1%; Score 36; DB 7; Length 856;
Best Local Similarity 77.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILEENKELE 9
    ||||| |
DB 170 ILEENNEIE 178

RESULT 30
US-11-087-099-3432
; Sequence 3432, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3432
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(186)
; OTHER INFORMATION: unsure at all xaa.locations
US-11-087-099-3432

Query Match      56.5%; Score 35; DB 7; Length 186;
Best Local Similarity 70.0%; Pred. No. 82;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EENKELENAL 12
    ||||| |
DB 33 EENKRPESAL 42

RESULT 31
US-11-087-099-7178
; Sequence 7178, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7178

; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 116
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-054-281-116

Query Match      56.5%; Score 35; DB 7; Length 247;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
    ||||| |
DB 65 LEKNFEMENKLE 76

RESULT 32
US-11-219-282-29
; Sequence 29, Application US/11219282
; Publication No. US2006001892A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: DuBose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-A
; CURRENT APPLICATION NUMBER: US/11/219,282
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US/10/094,944
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent version 3.1
; SEQ ID NO 29
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-219-282-29

Query Match      56.5%; Score 35; DB 7; Length 255;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LEENKELENAL 12
    ||||| |
DB 235 LOENRSMESAL 245

RESULT 33
US-10-714-887-274
; Sequence 274, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
```

```
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 274
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1634 polypeptide; paralogous to G2701
US-10-714-887-274

Query Match          56.5%; Score 35; DB 6; Length 277;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 EENKELENAL 12
      ||||: |||
Db      34 EENKFERAL 43

RESULT 34
US-11-087-099-1217
; Sequence 1217, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1217
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-1217

Query Match          56.5%; Score 35; DB 7; Length 277;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 EENKELENAL 12
      ||||: |||
Db      34 EENKFERAL 43

RESULT 35
US-10-467-657-3852
; Sequence 3852, Application US/10467657
; Publication No. US20050260591A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
```

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3852
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3852

Query Match          56.5%; Score 35; DB 6; Length 338;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 ILEENKELENAL 13
      ||||: ||||
Db      304 ILEKSEPLEDSLK 316

RESULT 36
US-10-793-626-2250
; Sequence 2250, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2250
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2250

Query Match          56.5%; Score 35; DB 6; Length 446;
Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      2 LEENKELENAL 12
      ||||: |||
Db      401 MRENKSIKNAL 411

RESULT 37
US-11-098-686-10757
; Sequence 10757, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10757
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
```

US-11-098-686-10757

Query Match 56.5%; Score 35; DB 7; Length 481;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LEENKELENALK 13
|:|:|:|:|:|
DB 352 LDNEELRNEWK 363

RESULT 38

US-11-191-374-3
; Sequence 3, Application US/11191374
; Publication No. US20050260673A1
; GENERAL INFORMATION:
; APPLICANT: Hresko, Michelle Coutu
; APPLICANT: McLaird, Merry B.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Frevert, Anita M.
; APPLICANT: Chiapelli, Brandi
; APPLICANT: Baublite, Catherine
; APPLICANT: Kloek, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: 12557-015001
; CURRENT APPLICATION NUMBER: US/11/191,374
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US/10/771,708
; PRIOR FILING DATE: 2004-02-04
; PRIOR APPLICATION NUMBER: US 60/444,771
; PRIOR FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Strongyloides stercoralis
US-11-191-374-3

Query Match 56.5%; Score 35; DB 7; Length 1122;
Best Local Similarity 58.3%; Pred. No. 6.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENAL 12
|:|:|:|:|:|
DB 215 IAVENKQIENEL 226

RESULT 39

US-11-191-375-3
; Sequence 3, Application US/11191375
; Publication No. US20050260674A1
; GENERAL INFORMATION:
; APPLICANT: Hresko, Michelle Coutu
; APPLICANT: McLaird, Merry B.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Frevert, Anita M.
; APPLICANT: Chiapelli, Brandi
; APPLICANT: Baublite, Catherine
; APPLICANT: Kloek, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: 12557-015001
; CURRENT APPLICATION NUMBER: US/11/191,375
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US/10/771,708

; PRIOR FILING DATE: 2004-02-04
; PRIOR APPLICATION NUMBER: US 60/444,771
; PRIOR FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Strongyloides stercoralis
US-11-191-375-3

Query Match 56.5%; Score 35; DB 7; Length 1122;
Best Local Similarity 58.3%; Pred. No. 6.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENAL 12
|:|:|:|:|:|
DB 215 IAVENKQIENEL 226

RESULT 40

US-11-191-588-3
; Sequence 3, Application US/11191588
; Publication No. US20050282222A1
; GENERAL INFORMATION:
; APPLICANT: Hresko, Michelle Coutu
; APPLICANT: McLaird, Merry B.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Frevert, Anita M.
; APPLICANT: Chiapelli, Brandi
; APPLICANT: Baublite, Catherine
; APPLICANT: Kloek, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: 12557-015001
; CURRENT APPLICATION NUMBER: US/11/191,588
; CURRENT FILING DATE: 2005-07-28
; PRIOR APPLICATION NUMBER: US/10/771,708
; PRIOR FILING DATE: 2004-02-04
; PRIOR APPLICATION NUMBER: US 60/444,771
; PRIOR FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Strongyloides stercoralis
US-11-191-588-3

Query Match 56.5%; Score 35; DB 7; Length 1122;
Best Local Similarity 58.3%; Pred. No. 6.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILEENKELENAL 12
|:|:|:|:|:|
DB 215 IAVENKQIENEL 226

RESULT 41

US-11-072-175-192
; Sequence 132, Application US/11072175
; Publication No. US20060029944A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273A CIP
; CURRENT APPLICATION NUMBER: US/11/072,175
; CURRENT FILING DATE: 2005-03-05

; PRIOR APPLICATION NUMBER: US 60/406,385
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 10/648,593
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 571
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 192
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-175-192

Query Match 56.5%; Score 35; DB 7; Length 1261;
Best Local Similarity 69.2%; Pred. No. 7.7e+02;
Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy 3 EENK--ELENAK 13
Db 314 EQNMLEAENALK 326
:|:|:|:|:|:|

RESULT 42
US-11-196-400-3
; Sequence 3, Application US/11196400
; Publication No. US20050287166A1
; GENERAL INFORMATION:
; APPLICANT: DRUIIHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773US01V
; CURRENT APPLICATION NUMBER: US/11/196,400
; CURRENT FILING DATE: 2005-08-04
; PRIOR APPLICATION NUMBER: US/09/742,096
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,462
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-11-196-400-3

Query Match 56.5%; Score 35; DB 7; Length 1786;
Best Local Similarity 77.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILEENKE 9
Db 1355 ILEEKKEIE 1363
|:|:|:|:|:|

RESULT 43
US-10-793-626-1080
; Sequence 1080, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1080
; LENGTH: 582

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1080

Query Match 55.6%; Score 34.5; DB 6; Length 582;
Best Local Similarity 56.2%; Pred. No. 3.8e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 1 ILEE--NKELENALK 13
Db 417 VLBEELGNKEANELK 432
:|:|:|:|:|:|

RESULT 44
US-10-467-657-7832
; Sequence 7832, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7832
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7832

Query Match 54.8%; Score 34; DB 6; Length 63;
Best Local Similarity 58.3%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LEENKELENALK 13
Db 34 LEERHDLERAVK 45
|:|:|:|:|:|

RESULT 45
US-11-096-568A-32955
; Sequence 32955, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32955
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(125)
; OTHER INFORMATION: Ceres Seq. ID no. 13601567
US-11-096-568A-32955

Query Match 54.8%; Score 34; DB 7; Length 125;
Best Local Similarity 46.2%; Pred. No. 75;

```
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 101 LLEDNQNLNRQLE 113
RESULT 46
US-11-096-568A-32954
; Sequence 32954, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32954
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(206)
; OTHER INFORMATION: Ceres Seq. ID no. 13601566
US-11-096-568A-32954
Query Match 54.8%; Score 34; DB 7; Length 206;
Best Local Similarity 46.2%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 182 LLEDNQNLNRQLE 194
RESULT 47
US-11-096-568A-32953
; Sequence 32953, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 32953
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(218)
; OTHER INFORMATION: Ceres Seq. ID no. 13601565
US-11-096-568A-32953
Query Match 54.8%; Score 34; DB 7; Length 218;
Best Local Similarity 46.2%; Pred. No. 1.4e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 ILEENKELENALK 13
Db 194 LLEDNQNLNRQLE 206
RESULT 48
US-11-087-099-11150
; Sequence 11150, Application US/11087099
; Publication No. US20060041961A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11150
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-11-087-099-11150
Query Match 54.8%; Score 34; DB 7; Length 259;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 ILEENKELENALK 13
Db 25 LEKNFEVENKLE 36
RESULT 49
US-11-087-099-10418
; Sequence 10418, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10418
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-10418
Query Match 54.8%; Score 34; DB 7; Length 267;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 EENKELENAL 12
Db 33 EENKRFEXAL 42
RESULT 50
US-11-196-475-92
; Sequence 92, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-92

Query Match          54.8%; Score 34; DB 7; Length 273;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
DB      266 KELKNALK 273
|||||

RESULT 51
US-11-196-475-98
; Sequence 98, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-98

Query Match          54.8%; Score 34; DB 7; Length 273;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
DB      266 KELKNALK 273
|||||

RESULT 52
US-11-196-475-101
; Sequence 101, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
```

```
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-101

Query Match          54.8%; Score 34; DB 7; Length 273;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
DB      266 KELKNALK 273
|||||

RESULT 53
US-11-196-475-142
; Sequence 142, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-142

Query Match          54.8%; Score 34; DB 7; Length 273;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
DB      266 KELKNALK 273
|||||
```

```
; ORGANISM: Borrelia burgdorferi
US-11-196-475-9

Query Match      54.8%; Score 34; DB 7; Length 274;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
      |||:||||
Db      267 KELKNALK 274

RESULT 56
US-11-087-099-7007
; Sequence 7007, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7007
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-087-099-7007

Query Match      54.8%; Score 34; DB 7; Length 288;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 ENKELENAL 12
      |||:||||
Db      38 ENKAFENAL 46

RESULT 57
US-10-793-626-66
; Sequence 66, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-66

Query Match      54.8%; Score 34; DB 6; Length 297;
Best Local Similarity 38.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY      1 ILEENKELENALK 13
      ::::|::|
Db      238 LIKDNPEIDKALK 250

RESULT 58
US-10-793-626-746
; Sequence 746, Application US/10793626
```

```
US-11-196-475-144
; Sequence 144, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-144

Query Match      54.8%; Score 34; DB 7; Length 273;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      6 KELENALK 13
      |||:||||
Db      266 KELKNALK 273

RESULT 55
US-11-196-475-9
; Sequence 9, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 274
; TYPE: PRT
```

; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: *2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 746
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-746

Query Match 54.8%; Score 34; DB 6; Length 297;
Best Local Similarity 38.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ILEENKELENALK 13
Db 238 LIKNPEIDKALK 250
:::|:|:|

RESULT 59
US-11-087-099-9768
; Sequence 9768, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9768
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Antirrhinum majus
US-11-087-099-9768

Query Match 54.8%; Score 34; DB 7; Length 307;
Best Local Similarity 77.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 ENKELENAL 12
Db 30 ENKAFENAL 38
||| |||

RESULT 60
US-11-072-512-3424
; Sequence 3424, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3424
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3424

Query Match 54.8%; Score 34; DB 7; Length 341;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILEENKELEN 10
Db 149 LLEPDKEMEN 158
:|:|:|:|

Search completed: March 11, 2006, 12:33:01
Job time : 14.9737 secs